

OPI DATE 10/02/99 APPLN. ID 83252/98
AOJP DATE 01/04/99 PCT NUMBER PCT/AU98/00564



AU9883252

AL

INT

(51) International Patent Classification 6 : C07K 14/435, C07H 21/04, A61K 38/17, C12N 15/12, C12P 21/02		A1	(11) International Publication Number: WO 99/03886 (43) International Publication Date: 28 January 1999 (28.01.99)
(21) International Application Number: PCT/AU98/00564 (22) International Filing Date: 17 July 1998 (17.07.98) (30) Priority Data: PO 8117 18 July 1997 (18.07.97) AU (71) Applicant (for all designated States except US): THE UNIVERSITY OF SYDNEY [AU/AU]; Parramatta Road, Sydney, NSW 2006 (AU). (72) Inventor; and (75) Inventor/Applicant (for US only): WEISS, Anthony, Steven [AU/AU]; 235 Rainbow Street, Randwick, NSW 2031 (AU). (74) Agent: GRIFFITH HACK; G.P.O. Box 4164, Sydney, NSW 2001 (AU).		(81) Designated States: AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, CA, CH, CN, CU, CZ, DE, DK, EE, ES, FI, GB, GE, GH, GM, HR, HU, ID, IL, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MD, MG, MK, MN, MW, MX, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, UA, UG, US, UZ, VN, YU, ZW, ARIPO patent (GH, GM, KE, LS, MW, SD, SZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG). Published With international search report.	
(54) Title: TROPOELASTIN DERIVATIVES			
(57) Abstract The invention relates to derivatives of tropoelastin and variants of those derivatives. The invention further provides expression products and hybrid molecules of the derivatives and variants of the invention. The invention further provides methods for the production of the derivatives, variants, expression products and hybrid molecules. Further provided are formulations, cross-linked structures and implants comprising the derivatives, variants, expression products and hybrid molecules of the invention. Further provided are uses of the derivatives, variants, expression products and hybrid molecules of the invention.			

TROPOELASTIN DERIVATIVES

TECHNICAL FIELD

The present invention relates to derivatives of human
5 tropoelastin and variants thereof, to genetic constructs
encoding the amino acid sequences of the derivatives and
variants and to uses of the derivatives and variants. In
particular, the derivatives of the present invention have
elastin-like properties or macro-molecular binding
10 properties.

BACKGROUND ART

There are various forms of tropoelastin that
typically appear to consist of two types of alternating
15 domains: those rich in hydrophobic amino acids
(responsible for the elastic properties) and those rich in
lysine residues (responsible for cross-link formation).
Hydrophobic and cross-linking domains are encoded in
separate exons (Indik et al 1987).

20 The 26 A region of human tropoelastin is unique
amongst tropoelastin domains in that, due to the absence
of lysine, this region does not participate in elastin
cross-link formation. Furthermore, this region is a
serine-rich domain and lacks hydrophobic stretches,
25 indicating that it is unlikely to contribute to the
elasticity of tropoelastin. There is otherwise limited
information on the structure and functional relationships
of the 26 A region (Bedell-Hogan et al., 1993).

The gene for tropoelastin is believed to be present
30 as a single copy in the mammalian genome, and is expressed
in the form of multiple transcripts, distinguished by
alternative splicing of the pre-mRNA (Indik et al, 1990;
Oliver et al, 1987). Modest expression of a natural human
tropoelastin sequence has been achieved by Indik et al
35 (1990) using cDNA, providing free polypeptide which
unfortunately was unstable.

Expression of substantial amounts of human
tropoelastin using synthetic polynucleotides is reported

in WO94 14958. In particular, a construct SHL providing substantial amounts of full length human tropoelastin is described.

5

DESCRIPTION OF THE INVENTION

In the specification and claims, "derivatives of human tropoelastin" or "tropoelastin derivatives" means novel peptides, polypeptides or proteins which contain amino acid sequences derived from the native amino acid sequences of human tropoelastin molecules. The amino acid sequences of the derivatives of human tropoelastin may be derived from any of the amino acid sequences of the isoforms of human tropoelastin. Derivatives of human tropoelastin are distinguished from human tropoelastin molecules in that the amino acid sequences of derivatives are altered with respect to native tropoelastin sequences by substitution, addition or deletion of residues, or a combination of these alterations, in derivative amino acid sequences.

20 In a first aspect, the present invention provides derivatives of human tropoelastin which have elastin-like properties. Elastin-like properties are a combination of elastic properties, including the phenomenon of recoil following molecular distention under appropriate conditions, and the ability to be cross-linked to other elastin molecules and/or other elastin-like molecules.

25 In a second aspect, the present invention provides derivatives of human tropoelastin which have macro-molecular binding properties including the ability to bind glycosaminoglycans.

30 In a third aspect, the present invention provides derivatives of human tropoelastin which have elastin-like properties and macro-molecular binding properties.

The present invention further provides amino acid sequence variants of the derivatives of the invention. In the specification and claims "variants" means amino acid sequences which retain the properties of the corresponding derivative of human tropoelastin, for example, elastin-

- 3 -

like properties or macro-molecular binding properties, or a combination of elastin-like properties and macro-molecular binding properties, and have an amino acid sequence which is homologous with the amino acid sequence of the corresponding derivative. For the purposes of this description, "homology" between the amino acid sequence of a particular derivative of human tropoelastin and another amino acid sequence connotes a likeness short of identity, indicative of a derivation of one sequence from the other. In particular, an amino acid sequence is homologous to a derivative of human tropoelastin if the alignment of that amino acid sequence with the sequence of the derivative of human tropoelastin reveals a similarity of about 65% over any 20 amino acid stretch or over any repetitive element of the molecules shorter than 20 amino acids in length. Such a sequence comparison can be performed via known algorithms, such as that of Lipman and Pearson (1985). Similarity is observed between amino acids where those amino acids have a side chain which confers a similar chemical property in the same chemical environment. For example, threonine and serine are similar amino acids; aspartic acid and glutamic acid are similar amino acids; valine, leucine and isoleucine are similar amino acids etc. Thus, an amino acid sequence may be considered homologous with the amino acid sequence of a human tropoelastin derivative because the alignment of those sequences reveals a similarity of 65%, although at each amino acid position in the aligned sequences, none of the residues are identical.

Inasmuch as the present invention provides derivatives of human tropoelastin and amino acid sequence variants of those derivatives, the invention thus extends to amino acid sequence variants incorporating amino acid sequences of non-human tropoelastins. Amino acid sequence variants which are non-human tropoelastin derivatives, or are based all, or in part, on non-human tropoelastin derivatives retain properties of the corresponding

elastin like properties, or elastin like properties, or elastin like properties and macro-molecular binding properties, and have an amino acid sequence which is homologous with the amino acid sequence of the corresponding human derivative. The variants of the invention also include variants of the non-human tropoelastin derivatives, or of derivatives based on the non-human tropoelastin derivatives.

"Homology" between the amino acid sequence of a particular derivative of non-human tropoelastin and another amino acid sequence connotes a likeness short of identity, indicative of a derivation of one sequence from the other. In particular, an amino acid sequence is homologous to a derivative of non-human tropoelastin if the alignment of that amino acid sequence with the sequence of the derivative of non-human tropoelastin reveals a similarity of about 65% over any 20 amino acid stretch or over any repetitive element of the molecules shorter than 20 amino acid in length. The skilled addressee will understand that species that are substantially phylogenetically related to humans express tropoelastin molecules which consist of amino acid sequences with homology to human tropoelastin amino acid sequences. Indeed, amino acid sequences of non-human tropoelastins have been determined, including the amino acid sequences of chick tropoelastin, bovine tropoelastin and rat tropoelastin (Bressan et al. 1987, Raju et al. 1987, Pierce et al. 1992) and over multiple regions, these are homologous with the human tropoelastin amino acid sequences. The skilled addressee will recognise therefore, that derivatives of human tropoelastin and amino acid sequence variants of those derivatives will necessarily encompass corresponding tropoelastin amino acid sequences from these and other non-human species.

The present invention provides a tropoelastin derivative comprising the amino acid sequence of SHEL8modified (SEQ ID NO:5). The amino acid sequence of

- 5 -

SHELδmodified and the alignment of that amino acid sequence with the human tropoelastin sequence is shown in Figure 5.

The invention also provides an amino acid sequence variant of the derivative comprising the amino acid sequence of SHELδmodified.

The invention also provides a polynucleotide encoding a tropoelastin derivative comprising the amino acid sequence of SHELδmodified. The nucleotide sequence encoding SHELδmodified is shown in Figure 3 (SEQ ID NO: 4). Preferably the polynucleotide comprises the nucleotide sequence which corresponds to SHELδmodified shown in Figure 3.

The invention also provides a polynucleotide encoding an amino acid sequence variant of the derivative SHELδmodified.

The present invention further provides a synthetic polynucleotide encoding a tropoelastin derivative comprising the amino acid sequence of SHELδ26A (SEQ ID NO:3). A synthetic polynucleotide is a molecule which comprises a nucleotide sequence that contains silent mutations with respect to the corresponding native polynucleotide molecule. The silent mutations enhance the expression of the synthetic polynucleotide. The amino acid sequence of SHELδ26A and the alignment of that amino acid sequence with the human tropoelastin sequence is shown in Figure 2. The SHELδ26A derivative excludes the SHEL coding sequence corresponding to exon 26A. Preferably the synthetic polynucleotide comprises the sequence shown in Figure 1 (SEQ ID NO:1) from nucleotide position 1 to 1676 contiguous with nucleotide position 1775 to 2210.

The invention also provides a polynucleotide encoding an amino acid sequence variant of the derivative SHELδ26A.

The invention also provides an amino acid sequence

- 6 -

variant of the derivative comprising the amino acid sequence of SHEL026A.

The present inventor has, for the first time, shown that the region encoded by exon 26A (peptide 26A) of the tropoelastin gene binds glycosaminoglycans (GAGs) (Figure 6A and B). GAGs are macro-molecules particularly associated with the extracellular environment. These molecules play an important role in the architecture and mechanical properties of connective tissues and mediate interactions with and availability of other molecules.

Thus, the present invention provides a tropoelastin derivative comprising the amino acid sequence of peptide 26A. Peptide 26A has the amino acid sequence:

GADEGVRRSLSPELREGDPSSSQHLPSTPSSPEV (SEQ ID NO: 12) or
GADEGVRRSLSPELREGDPSSSQHLPSTPSSPEF (SEQ ID NO: 13).

claim 77
claims
57, 58, 64, 65

The present invention also provides an amino acid sequence variant of the derivative comprising the amino acid sequence of peptide 26A.

The invention also provides a polynucleotide encoding a tropoelastin derivative comprising the amino acid sequence of peptide 26A. Preferably the polynucleotide comprises the nucleotide sequence shown in Figure 1 (SEQ ID NO: 1) from nucleotide position 1687 to 1778. Preferably the 3' terminal codon is GTT (which encodes V) or TTT (which encodes F).

The invention also provides a polynucleotide encoding an amino acid sequence variant of the derivative comprising the amino acid sequence of peptide 26A.

In appreciating the GAG binding property of peptide 26A, the present inventor envisages the generation of novel subsets of hybrid molecules, comprising biological polymers which are linked to peptide 26A, wherein the peptide 26A imparts GAG binding activity to the polymer. In particular, the present inventor has recognised that the deletion or insertion of the peptide 26A amino acid sequence, or a variant of that amino acid sequence will alter GAG binding activity. Thus, the present invention relates to tropoelastin derivatives in which full length

or partial length tropoelastin molecules have been modified by the addition of one or more extra 26A regions to enhance interactions with GAGs. Moreover, the invention relates to site directed modification of the amino acid sequence of peptide 26A so as to generate variants of the peptide 26A amino acid sequence which have altered affinity or altered specificity for GAGs. Tropoelastin derivatives or variants of the derivatives which contain altered GAG binding activity may be uncross-linked or cross-linked.

In another aspect, the invention provides a hybrid molecule. In the specification and claims, "hybrid molecule" means a molecule comprising a biological polymer which is linked to a tropoelastin derivative comprising the amino acid sequence of peptide 26A or an amino acid sequence variant of a derivative comprising the amino acid sequence of peptide 26A. Preferably the biological polymer is a protein. More preferably the protein is selected from the group consisting of growth factors, cytokines and antibodies. Alternatively the biological polymer is selected from the group consisting of lipids, sugars or nucleic acids.

In one embodiment, and where the biological polymer is a protein, the hybrid molecule is produced by recombinant DNA techniques, including for example the construction of a nucleotide sequence which encodes the biological polymer and the tropoelastin derivative comprising the amino acid sequence of peptide 26A, or the amino acid sequence variant of a derivative comprising the amino acid sequence of peptide 26 A, in a single open reading frame. Alternatively, the hybrid molecule may be produced synthetically by solid phase peptide synthesis, including, for example the methods of synthesis disclosed in Merrifield (1963) or Knorr et al. (1989). Examples of peptide synthesis also include the synthesis methods used by peptide synthesizers of Perkin Elmer/Applied Biosystems, CA, US.

In another aspect, the invention provides a

- 5 -

polynucleotide sequence encoding a hybrid molecule of the invention.

In another aspect, the invention provides a hybrid molecule which comprises a synthetic polymer which is
5 linked in a tropoelastin derivative comprising the amino acid sequence of peptide 26A, or an amino acid sequence variant of the derivative comprising the amino acid sequence of peptide 26A.

The invention further provides a method of imparting
10 or enhancing GAG binding activity to a biological polymer comprising the step of linking a tropoelastin derivative comprising the amino acid sequence of peptide 26A, or an amino acid sequence variant of peptide 26A with the biological polymer. Preferably the biological polymer is
15 a protein.

The invention further provides a method of deleting or reducing GAG binding activity from a biological polymer comprising the step of deleting a tropoelastin derivative comprising the amino acid sequence of peptide 26A, or an
20 amino acid sequence variant of peptide 26A from the biological polymer. Preferably the biological polymer is a protein.

The present invention also provides a tropoelastin derivative comprising the amino acid sequence of
25 SHELgamma. SHELgamma has the amino acid sequence:
SANGALVGLGVPGLVGAGVPGFGAGADEGVRRSLSPELREGDPSSSQHLPSTPSSPR
VPGALAAAKAAKYGAAPVPGVLGGGLGALGGVGIPGGVVGAGPAAAAAAKAAAKAAQFG
LVGAAGLGGGLGVGGLGVPGVGGLGGIPPAKAAKYGAAGLGGVLGGAGQFPLGGVA
ARPGFGLSPIFPGGACLGKACGRKRK (SEQ ID NO: 9).

30 The invention also provides an amino acid sequence variant of the derivative comprising the amino acid sequence of SHELgamma.

The invention also provides a polynucleotide encoding a tropoelastin derivative, the derivative comprising the
35 amino acid sequence of SHELgamma. The nucleotide sequence of the polynucleotide SHELgamma (SEQ ID NO: 8) is shown in Figure 8. In this nucleotide sequence, the first 9 codons from nucleotide position 948 to 974 are derived

- 5 -

from the glutathione S-transferase GST fusion nucleotide sequence. Preferably the polynucleotide comprises the nucleotide sequence shown in Figure 5. More preferably the polynucleotide comprises the nucleotide sequence shown in Figure 5 from nucleotide sequence position 975 to 1547.

The invention also provides a polynucleotide encoding an amino acid sequence variant of the derivative comprising the amino acid sequence of SHELgamma.

The present invention also provides a polynucleotide encoding a tropoelastin derivative, the derivative comprising the amino acid sequence of SHELgamma excluding exon 26A. The nucleotide sequence of the polynucleotide SHELgamma excluding exon 26A (SEQ ID NO: 6) is shown in Figure 7. In this nucleotide sequence, the first 5 codons from nucleotide position 948 to 962 are derived from the GST nucleotide sequence. SHELgamma excluding exon 26A has the following amino acid sequence:

VPGALAAAKAAKYGFAYPGVLGGGLGALGGVGIPOGGVTCASPAAAAAAKAAAKAAQFG
LVGAAGLGGGLGVGGLGVPGVGGGLGCIPPAAAAKAAKYGAAGLGGVLCGAGQFFLGVA
ARPGFGLSPIFFGGACLGKACGRKRK (SEQ ID NO: 7).

Preferably the polynucleotide comprises the nucleotide sequence shown in SEQ ID NO: 6. More preferably the polynucleotide comprises the nucleotide sequence shown in SEQ ID NO: 6 from nucleotide sequence position 15 to 441.

The invention also provides a polynucleotide encoding an amino acid sequence variant of the derivative comprising the amino acid sequence of SHELgamma excluding exon 26A.

The invention also provides a tropoelastin derivative comprising the amino acid sequence of SHELgamma excluding exon 26A.

The invention also provides an amino acid sequence variant of the derivative comprising SHELgamma excluding exon 26A.

The derivatives of the invention based on SHELgamma can also be produced by in vitro biochemical cleavage of tropoelastin products such as SHEL, so as to release a carboxy-terminal fragment. The carboxy-terminal fragment

claims
57, 66, 6-
68
629, 26
23

- 11 -

may be purified by reverse phase HPLC.

The present invention also provides a tropoelastin derivative comprising the amino acid sequence of SHEL31-36. SHEL31-36 has the following amino acid sequence:

5 GIPFAAAAKAAKYGAAGLGGVLGGAGQFFLGGVAARPGFGLSPIFFGGACLGKACG-
RKRR (SEQ ID NO: 10).

SHEL31-36 retains a crosslinking domain. As a consequence of its elastin-like properties, it is envisaged that this and related tropoelastin derivatives
10 can be used to interfere with tropoelastin deposition and formation of unaltered elastic fibre.

The invention also provides an amino acid sequence variant of the derivative comprising the amino acid sequence of SHEL31-36.

15 The invention also provides a polynucleotide encoding a tropoelastin derivative, the derivative comprising the amino acid sequence of SHEL31-36. Preferably the polynucleotide comprises the nucleotide sequence shown in Figure 1 (SEQ ID NO:1) from nucleotide position 2012 to
20 2210.

The invention also provides a polynucleotide encoding an amino acid variant of the derivative comprising the amino acid sequence of SHEL31-36.

25 The present invention also provides a tropoelastin derivative, comprising the amino acid sequence of SHEL32-36. SHEL32-36 has the following amino acid sequence:
GAAGLGGVLGGAGQFPLOGVAARPGFGLSPIFFGGACLGKACGRKRR (SEQ ID NO: 11).

30 The invention also provides an amino acid sequence variant of the derivative comprising the amino acid sequence of SHEL32-36.

The invention also provides a polynucleotide encoding a tropoelastin derivative, the derivative comprising the amino acid sequence of SHEL32-36. Preferably the
35 polynucleotide comprises the nucleotide sequence shown in Figure 1 (SEQ ID NO: 1) from nucleotide position 2061 to 2210.

The present invention also provides a polynucleotide

- 11 -

to 1076 contiguous with 1776 to 2211.

The present invention also provides a polynucleotide encoding an amino acid sequence variant of the derivative comprising the amino acid sequence of SH2106-36.

5 In another aspect the present invention provides a formulation comprising a tropoelastin derivative, a variant of the derivative or a hybrid molecule of the invention together with a carrier or diluent.

10 Formulations of the derivatives, variants or hybrid molecules of the invention can be prepared in accordance with standard techniques appropriate to the field in which they are to be used.

The polynucleotides and synthetic polynucleotides of the invention can be provided in association with other
15 polynucleotide sequences including 5' and 3' untranslated sequences, and 5' upstream and 3' downstream transcriptional regulatory sequences. The polynucleotides and synthetic polynucleotides may be provided as a recombinant DNA molecule including plasmid DNA.

20 The polynucleotides and synthetic polynucleotides of the invention can be prepared using the techniques of chemical synthesis or recombinant DNA technology, or by a combination of both techniques.

In a further aspect the invention provides a vector
25 comprising a polynucleotide or synthetic polynucleotide encoding a tropoelastin derivative, a variant of the derivative or a hybrid molecule of the invention.

Vectors useful in this invention include plasmids, phages and phagemids. The polynucleotides and synthetic
30 polynucleotides of the present invention can also be used in integrative expression systems or lytic or comparable expression systems.

Suitable vectors will generally contain origins of replication and control sequences which are derived from
35 species compatible with the intended expression host. Typically these vectors include a promoter located upstream from the polynucleotide, together with a ribosome binding site if intended for prokaryotic expression and a

- 13 -

phenotypic selection gene such as one conferring antibiotic resistance or supplying an auxotrophic requirement. For production vectors, vectors which provide for enhanced stability through partitioning may be
5 chosen. Where integrative vectors are used it is not necessary for the vector to have an origin of replication. Lytic and other comparable expression systems do not need to have those functions required for maintenance of vectors in hosts.

10 For *E. coli* typical vectors include pBR322, pBluescript II SK⁺, pGEX-2T, pTrc99A, pET series vectors, particularly pET3d, (Studier et al., 1990) and derivatives of these vectors. Derivatives include those plasmids with a modified protease recognition sequence to facilitate
15 purification of a protein domain.

In another aspect the invention provides a cell capable of expressing a polynucleotide or a synthetic polynucleotide which encodes a derivative or variant of the invention, or a polynucleotide which encodes a hybrid
20 molecule of the invention.

A preferred expression system is an *E. coli* expression system. However, the invention includes within its scope the use of other hosts capable of expressing protein from the polynucleotides designed for use in *E.*
25 *coli*. The invention also includes the use of polynucleotides and synthetic polynucleotides suitable for use in other expression systems such as other microbial expression systems. These other expression systems include yeast, and bacterial expression systems, insect
30 cell expression systems, and expression systems involving other eukaryotic cell lines or whole organisms.

Examples of *E. coli* hosts include *E. coli* B strain derivatives (Studier et al, 1990), and K-strain derivatives such as NM522 (Gough and Murray, 1983) and
35 XL1-Blue (Bullock et al, 1987).

In a further aspect the present invention provides an expression product. In the specification and claims, expression product means a derivative or variant of the

invention expressed by a cell containing a polynucleotide or a synthetic polynucleotide encoding a derivative or variant of the invention.

The expression products of the invention may be fused
5 expression products which include all or part of a protein encoded by the vector in peptide linkage with the derivative or variant. They may also include, for example, an N-terminal methionine or other additional residues which do not permanently impair the elastin-like,
10 or macro-molecular binding properties of the product.

Typically the fusion is to the N-terminus of the expression product. An example of a suitable protein is to the C-terminus of glutathione S-transferase. The fused protein sequence may be chosen in order to cause the
15 expression product to be secreted or expressed as a cell surface protein to simplify purification or expressed as a cytoplasmic protein.

The expressed fusion products may subsequently be treated to remove the fused protein sequences to provide
20 free tropoelastin derivative or variant. Treatment is typically through protease treatment or, in the case of secretion, removal is effected by endogenous host secretion machinery. An example of this is secretion by yeasts.

25 Non-fused systems include the introduction of or use of a pre-existing methionine codon. An example of this is the use of pET3a or pET3d in *E. coli*.

In another aspect the invention provides a polynucleotide encoding an expression product of the
30 invention.

In another aspect the present invention provides a formulation comprising an expression product of the invention together with a carrier or diluent. The formulation of the expression product can be prepared in
35 accordance with standard techniques appropriate to the field in which they are to be used.

According to a further aspect of the present invention there is provided a method for producing a

- 15 -

tropoelastin derivative or a variant of the derivative comprising providing a vector containing a polynucleotide or a synthetic polynucleotide encoding the derivative or variant; introducing the vector into a suitable host cell; 5 maintaining the cell in conditions suitable for expression of the polynucleotide or synthetic polynucleotide and isolating the derivative or variant of the invention. The method can be applied to the production of the expression products and hybrid molecules (in which the hybrid 10 molecules comprise the peptide 26A or a variant thereof and a further amino acid sequence) of the invention, by providing a vector containing a polynucleotide encoding an expression product or a hybrid molecule; introducing the vector into a suitable host cell; maintaining the cell in 15 conditions suitable for expression of the polynucleotide and isolating the expression product or hybrid molecule.

In one embodiment, the polynucleotide or synthetic polynucleotide encoding the derivative, variant, expression product or hybrid molecule of the invention is 20 expressed in a host cell which is maintained in culture *in vitro*.

Alternatively, the polynucleotide or synthetic polynucleotide encoding the derivative, variant, expression product or hybrid molecule of the invention is 25 expressed in a host cell which is maintained *in vivo*. Thus, in another embodiment, the polynucleotide or synthetic polynucleotide encoding the derivative, variant, expression product or hybrid molecule of the invention is expressed in a transgenic animal. Methods for the 30 generation of transgenic animals are known in the art. Exemplary methods are described in Slack et al. 1991 and Janne et al. 1992.

The tropoelastin derivatives, variants of the derivatives, and hybrid molecules (in which the hybrid 35 molecules comprise the peptide 26A or a variant thereof and a further amino acid sequence) of the invention may be produced by solid phase peptide synthesis, including, for

disclosed in Merrifield

- 16 -

(1963) or Knorr et al (1989). Examples of peptide synthesis also include the synthesis methods used by peptide synthesisers of Perkin Elmer/Applied Biosystems, CA, US. As an alternative to cell synthesis from a polynucleotide or synthetic polynucleotide, the expression products of the invention may be produced by solid phase peptide synthesis.

In a further aspect the present invention provides an implant formed from at least one tropoelastin derivative and/or variant of the derivative of the invention. The implant may alternatively contain at least one expression product and/or at least one hybrid molecule of the invention.

The implants are formed into the required shape by cross-linking the tropoelastin derivative, variant of the derivative, expression product, or hybrid molecule of the invention, in a mould which conforms to the desired shape of the implant. Where the implant is required to be used in sheet form the tropoelastin derivative, variant of the derivative, expression product, or hybrid molecule of the invention can be cross-linked on a flat surface. Relevant methodologies are described in, for example, US Patent No. 4 474 851 and US Patent No. 5 250 516. The elastomeric materials may be exclusively prepared from one or more tropoelastin derivatives, variants of the derivative, expression products, or hybrid molecules of the invention or may be composites prepared from one or more of these constituents together with other materials.

The tropoelastin derivatives or variants of the derivatives can be cross linked to form elastin or elastin-like material or can be cross-linked in conjunction with other biological or synthetic molecules to form a composite material.

Thus in another aspect the invention provides a cross-linked complex which comprises at least one tropoelastin derivative of the invention and/or at least one variant of a derivative of the invention. The cross-linked complexes may additionally contain at least one

expression product and/or at least one hybrid molecule of the invention, which may be cross-linked to the at least one tropoelastin derivative and/or variant of the derivative of the invention.

5 The cross-linking of the tropoelastin derivatives, variants of the derivatives, hybrid molecules and expression products of the invention can be achieved by chemical oxidation of lysine side chains using processes such as ruthenium tetroxide mediated oxidation and quinone
10 mediated oxidation, or by using homobifunctional chemical cross-linking agents such as dithiobis (succinimidylpropionate), dimethyl adipimide or dimethyl pimelimide. Glutaraldehyde cross-linking is an important addition to this repertoire. Another alternative
15 is the cross-linking of lysine and glutamic side chains.

 The tropoelastin derivatives, variants of the derivatives, hybrid molecules and expression products of the invention may also be enzymatically cross-linked by methods including lysyl oxidase mediated oxidation or may
20 be cross-linked using gamma irradiation.

BRIEF DESCRIPTION OF THE DRAWINGS

 Figure 1: Nucleotide (SEQ ID NO: 1) and predicted amino acid (SEQ ID NO:2) sequences of synthetic human
25 tropoelastin SHEL. The upper (numbered) nucleotide sequence represents the coding strand.

 Figure 2: Alignment of SHEL (SEQ ID NO:2) (upper line) and SHEL δ 26A (SEQ ID NO: 3) amino acid sequences.

 Figure 3: Nucleotide (SEQ ID NO: 4) and predicted
30 amino acid (SEQ ID NO: 5) sequences of SHEL δ modified.

 Figure 4: Alignment of SHEL δ modified (SEQ ID NO: 4) (upper line) and SHEL (SEQ ID NO:1) nucleotide sequences.

 Figure 5: Alignment of SHEL δ modified (SEQ ID NO: 5) (lower line) and SHEL (SEQ ID NO: 1) amino acid
35 sequences.

 Figure 6A: HPLC elution profile of GST-exon 26A fusion protein tropoelastin derivative loaded in from

- 13 -

heparin sepharose. 6B: Binding of peptide 26A (SEQ ID NO: 12 and SEQ ID NO: 13) to glycosaminoglycans.

Figure 7: Nucleotide (SEQ ID NO: 6) and predicted amino acid (SEQ ID NO: 7) sequences of SHELgamma excluding
5 exon 26A.

Figure 8: Nucleotide (SEQ ID NO: 8) and predicted amino acid (SEQ ID NO: 9) sequences of SHELgamma.

BEST METHOD OF PERFORMING THE INVENTION

10 The recombinant and synthetic procedures used for the synthesis of the derivatives, variants, expression products and hybrid molecules of the invention are described in standard texts such as Sambrook et al (1989).

15 Tropoelastin nucleotide sequences may be modified so as to provide derivatives, variants, expression products or hybrid molecules, by conventional site-directed or random mutagenesis. The sequences may also be modified by oligonucleotide-directed mutagenesis, which comprises the following steps:

- 20 1. synthesis of an oligonucleotide with a sequence that contains the desired nucleotide substitution (mutation);
2. hybridising the oligonucleotide to a template comprising a structural sequence encoding
25 tropoelastin; and
3. using a DNA polymerase to extend the oligonucleotide as a primer.

Another approach which is particularly suited to situations where a synthetic polynucleotide encoding the
30 tropoelastin derivative is prepared from oligonucleotide blocks bounded by restriction sites, is cassette mutagenesis where entire restriction fragments are replaced.

35 Purification of the derivatives, variants, expression products or hybrid molecules of the invention is performed using standard techniques including HPLC. The actual sequence of steps in the purification of a particular derivative, variant, expression product or hybrid molecule

- 19 -

is limited by the environment from which the molecule is to be purified. By way of example, reference is made to the purification scheme disclosed in WO94 14955.

Formulations in accordance with the invention are
5 formulated in accordance with standard techniques.

The amount of derivative, variant, expression product or hybrid molecule that may be combined with a carrier or diluent to produce a single dosage will vary depending on the situation in which the formulation is to be used and
10 the particular mode of administration.

It will be understood also that specific doses for any particular host may be influenced by factors such as the age, sex, weight and general health of the host as well as the particular characteristics of the derivative,
15 variant, expression product or hybrid molecule of the invention being used, and how it is administered.

Injectable preparations, for example, sterile injectable aqueous or oleagenous suspensions may be formulated according to the known art using suitable
20 dispersing or wetting agents and suspending agents. The sterile injectable preparation may also be a sterile injectable solution or suspension in a non-toxic parenterally acceptable diluent or solvent. Among the acceptable vehicles or solvents that may be employed are
25 water, Ringer's solution, alcohols and isotonic sodium chloride solution. In addition, sterile, fixed oils are conventionally employed as a solvent or suspending medium. For this purpose any bland fixed oil may be employed including synthetic mono- or diglycerides. In addition,
30 fatty acids such as oleic acid and organic solvents find use in the preparation of injectables.

Routes of administration, dosages to be administered as well as frequency of administration are all factors which can be optimised using ordinary skill in the art.

35 In addition, the derivatives, variants, expression products and hybrid molecules of the invention may be prepared as topical preparations for instance as anti-wrinkle and hand lotions using standard techniques for the

- 23 -

preparation of such formulations. They may be prepared in aerosol form for, for instance, administration to a patient's lungs, or in the form of surgical implants, foods or industrial products by standard techniques.

5

SHEL

The preparation of SHEL is described in WO94 14958. It is directly expressed as a full length human protein with a calculated molecular weight of 64kDa. The full
10 nucleotide sequence and corresponding amino acid sequence of SHEL is shown in Figure 1. The preparation of pSHELF is described in WO94/14958.

pSHELF differs from the natural coding sequence(s) in a number of ways. As described in WO94 14958, the
15 untranslated regions present in the tropoelastin cDNA sequence were disregarded in designing the synthetic gene, and the nucleotides encoding the signal peptide were removed. Restriction endonuclease recognition sites were incorporated at regular intervals into the gene by
20 typically altering the third base of the relevant codons, thereby maintaining the primary sequence of the gene product. The facility for silent alteration of the coding sequence was also exploited to change the codon bias of the tropoelastin gene to that commonly found in highly
25 expressed *E.coli* genes. [Genetics Computer Group (GCG) package version 7-UNIX using Codon Frequency and Gen Run Data: ecohigh-cod]. Two additional stop codons were added to the 3'-end, and an ATG start codon comprising a novel NcoI site was appended to the 5'-end. *Bam* HI cloning sites
30 were engineered at both ends of the synthetic sequence. Since the gene contains no internal methionine residues, treatment of the newly-synthesized gene product (expressed directly or as a fusion with another gene) with cyanogen bromide would liberate a protein with the same or similar
35 sequence as one form of natural tropoelastin comprising 731 amino acids. Other forms of processing are envisaged, which may generate tropoelastin species of the same or different lengths

- 21 -

Two stop codons were added in order to allow the possible use of the construct in suppressor hosts, and also to avoid any potential depletion of termination (release) factors for translation.

5 As described in the following examples, the derivatives, pSHELF δ 26A, pSHELF δ modified, pSHELGamma, pSHEL31-36, pSHEL32-36 and pSHELGamma δ 26A were derived from the pSHELF nucleotide sequence. These particular derivatives, and indeed the derivatives, variants,
10 expression products and hybrid molecules of the invention can equally be derived from a native human or non-human tropoelastin nucleotide sequence.

Example 1: Construction of pSHELF δ 26A and pSHELF δ
15 modified

Mutagenesis was used with pSHELF to remove DNA corresponding to exon 26A. The sequence of the mutagenic primer was:

5'CGG GTT TCG GTG CTG TTC CGG GCG CCG TGG 3'

20 This flanked either side of exon 26A by 15bp resulting in its precise deletion. A second selection primer, which mutates a unique restriction site to another restriction site is normally used in the protocol but was not in this case since deletion of exon 26A also resulted
25 in the deletion of a unique restriction site, *Pml*I. The enzyme *Pml*I was used to treat the mutation reaction to linearise any unmutated parental plasmid and consequently to enrich for mutant plasmid. The reaction mixture was used to transform competent BMH17-18 *mutS* *E. coli*,
30 defective in mismatch repair, by electroporation and the entire transformed culture was grown overnight in LB+ampicillin. Mixed plasmid DNA, containing both mutated and parental plasmids, was isolated from the culture and the plasmid DNA was digested with *Pml*I to linearise the
35 parental plasmid. The plasmid DNA, now enriched for mutated plasmid, was used to transform *E. coli* HMS174 by electroporation and transformants selected on LB plates

containing 75µg/ml ampicillin.

Colonies were grown overnight and plasmid mini-preparations performed. Constructs were screened using PmlI and those which were insensitive to digestion were further screened by KpnI PstI double digestion. Candidate clones were sequenced to verify the sequence, named pSHELFδmodified.

Sequencing confirmed the region immediately surrounding the deletion was correct. PstI and BssHII restriction sites surrounding the correct region of pSHELFδmodified were used to remove the desired segment and re-insert it into the corresponding site of pSHELF. 6.5µg pSHELF and 7.5µg pSHELFδmodified were digested with BssHII, precipitated and digested with PstI. The appropriate three fragments were gel-purified and ligated. DNA was transformed into E. coli XL1-Blue and transformants selected on plates containing 75µg/ml ampicillin.

Plasmids were isolated by mini-preparations and screened using BglI digestion. A candidate clone was further analysed by restriction enzyme digestion and sequenced, and named pSHELF\$26A.

Example 2: Synthesis of Exon 26A

The region of SHEL corresponding to exon 26A was amplified by PCR, with primers modified to introduce an in-frame BamHI site upstream and a stop codon downstream at the 3' end. Two forms were generated: one terminating in valine (26AV) and the other terminating in phenylalanine (26AF). These forms are as follows:

GADEGVRRSLSPELREGDPSSSQHLPSTPSSPRV with properties:

Molecular weight = 3588.80

Residues = 34

Average Residue Weight = 105.553

Charge = -1

Isoelectric point = 5.71

- 13 -

and

GADEGVRRSLSPELREDDPSSSQHLFSTRSSPRF

A 26A coding region was expressed as a glutathione S-transferase (GST) fusion protein.

5

Example 3: Glycosaminoglycan-binding activity of Exon 26A

Ultrafiltration assay methodology was developed to examine and quantify interactions occurring in vitro between the 26A region and purified extracellular matrix glycosaminoglycans. GST26A fusion protein and tropoelastin were compared with GST and tropoelastin lacking exon 26A at physiologically relevant conditions of pH and ionic strength.

Experimental evidence supports the notion that peptide 26A (26AF and 26AV) binds GAGs. Immobilised heparin column binding shows that GST26A binds more tightly than does GST, and requires a higher sodium chloride concentration for elution (Figure 6B). Furthermore, GST26A fusion protein binds radioactive heparin with greater efficiencies than GST, and these can be compared with GAGs including chondroitin sulphates and keratin sulphates. An implication of this is that GAGs binding to tropoelastin can be adjusted based upon the content of 26A. Cross-linked tropoelastin would be expected to show differential binding to GAGs based on the relative amounts of SHEL vs. SHEL26A.

In summary, these studies reveal that the 26A region is a functional glycosaminoglycan binding domain, which functions in intact tropoelastin. It is also active when isolated as a fusion entity yet displays no detectable structure in the absence of bound GAG. Furthermore, panel competition studies indicate a preference for those GAGs found in close association with the elastic fibre in the extracellular matrix.

- 14 -

Example 4: Construction of pSHELgamma, pSHEL31-36,
pSHEL32-36 and pSHELgammaδ16A

pSHELgamma is derived from the pSHELgamma construct disclosed in WO94 14953. pSHEL31-36, pSHEL32-36 and
5 pSHELgammaδ16A were derived from pSHELgamma. pSHELgamma was modified by introducing an oligonucleotide linker at the XpnI site. This encoded a factor Xa cleavage site which could be utilised in subsequent constructs. PCR and
10 site directed mutagenesis was then used to generate further, shorter forms which provided fusions with GST. Constructs were DNA sequenced for verification. Induced protein was isolated as GST-fusion proteins, which were subsequently bound to glutathione agarose. Protease
15 cleavage was optional where fusion proteins were desired; otherwise the cleaved proteins and peptides were further purified by reverse phase HPLC.

INDUSTRIAL APPLICATION

The derivatives and expression products of the
20 invention are of use in inter alia the medical, pharmaceutical, veterinary and cosmetic fields.

- 25 -

REFERENCES

1. Indik D., Yeh H., Ornstein-Goldstein M., Sheppard F.,
Anderson N., Rosenbloom J.C., Peltonen L. and Rosenbloom
5 J. (1987) PNAS (USA) **84** 5661-5664
2. Indik D., Abrams W.R., Hucich U., Gibson C.W., Mechar
R.P. and Rosenbloom J. (1991) Arch. Biochem Biophys
10 **280** 80-86
3. Oliver L., Duvalle PA., Davidson J.M., Rosenbloom J.,
Mathew C.G., Betser A.J. and Boyd C.D. (1987)
Collagen Rel Res **7** 77-89
- 15 4. Sambrook J., Fritsch E.F., and Maniatis T. (1989)
Molecular cloning: a laboratory manual, second
edition Cold Spring Harbor Laboratory Press, Cold
Spring Harbor, New York
- 20 5. Bressan G.M., Argos P. and Stanley K.R. (1987)
Biochemistry **26** 1497-11503
6. Raju K. aand Anwar R.A. (1987) J. Biol Chem **262** 5755-
5762
- 25 7. Pierce R.A., Alatawi A., Deak S.B. & Boyd C.D. (1992)
Genomics **12** 651-658
8. Lipman and Pearson (1985) Science 227.1435.
- 30 9. Bedell-Hogan, D., Trackman, P., Abrams, W.,
Rosenbloom, J. and Kagan H. (1993) J. Biol. Chem.
268, 10345-10350
- 35 10. Studier, F. W., Rosenberg, A. H., Dunn, J. J. and
Dubendorff, J. W. (1990) Methods Enzymol. **185**, 60-89

1-19

12. Bullock, W. L., Fernandes, J. M. and Short, J. M.
(1997) *Biotechniques* **5** 376-379
- 5 13. Slack, J. L., Liska, D. J. and Bernstein P. (1991)
Mol. Cell Biol. **11**: 2166-2174
- 10 14. Janne, J., Hyttinen, J. M., Peura, T., Tolvanen, M.,
Alhonen, M. And Halmkyte M. (1992) *Ann. Med.* **24**:
273-280.
- 15 15. Merrifield, R.B., (1963) *J. Am. Chem. Soc.* **85**,
2149-2154.
16. Knorr R., Trzeciak, Samarth M., Gillesen, D. (1999),
Tetrahearon Letters **30**: 1927-1930

- 27 -

SEQUENCE LISTING

(1) GENERAL INFORMATION:

- (i) APPLICANT: WEISS, ANTHONY S
UNIVERSITY, SYDNEY
- (ii) TITLE OF INVENTION: TROPOELASTIN DERIVATIVES
- (iii) NUMBER OF SEQUENCES: 15
- (iv) CORRESPONDENCE ADDRESS:
 - (A) ADDRESSEE: GRIFFITH HACK
 - (B) STREET: 168 WALKER STREET
 - (C) CITY: NORTH SYDNEY
 - (D) STATE: NEW SOUTH WALES
 - (E) COUNTRY: AUSTRALIA
 - (F) ZIP: 2060
- (v) COMPUTER READABLE FORM:
 - (A) MEDIUM TYPE: Floppy disk
 - (B) COMPUTER: IBM PC compatible
 - (C) OPERATING SYSTEM: PC-DOS/MS-DOS
 - (D) SOFTWARE: PatentIn Release #1.0, Version #1.30
- (vi) CURRENT APPLICATION DATA:
 - (A) APPLICATION NUMBER: AU
 - (B) FILING DATE:
 - (C) CLASSIFICATION:
- (vii) PRIOR APPLICATION DATA:
 - (A) APPLICATION NUMBER: AU PO8117
 - (B) FILING DATE: 18-JUL-1997
- (viii) ATTORNEY/AGENT INFORMATION:
 - (A) NAME: GUMLEY, THOMAS P
 - (C) REFERENCE/DOCKET NUMBER: 04828ZK
- (ix) TELECOMMUNICATION INFORMATION:
 - (A) TELEPHONE: 61 2 9957 5944
 - (B) TELEFAX: 61 2 9957 6288
 - (C) TELEX: 26547

(2) INFORMATION FOR SEQ ID NO:1:

- 28 -

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 2210 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(iii) HYPOTHETICAL: YES

(iv) ANTI-SENSE: NO

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:1:

GATCCATGGG TGGCGTTCCG GGTGCTATCC CGGGTGGCGT TCCGGGTGGT GTATTCTACC	60
CAGGCGCGGG TCTGGGTGCA CTGGGCGGTG GTGCGCTGGG CCCGGGTGGT AAACCGCTGA	120
AACCGGTTCC AGGCGGTCTG GCAGGTGCTG GTCTGGGTGC AGGTCTGGGC GCGTTCCCGG	180
CGGTTACCTT CCCGGGTGCT CTGGTTCCGG GTGGCGTTGC AGACGCAGCT GCTGCGTACA	240
AAGCGGCAAA GGCAGGTGCG GGTCTGGGCG GGGTACCAGG TGTGCGCGT CTGGGTGTAT	300
CTGCTGGCGC AGTTGTTCCG CAGCCGGGTG CAGGTGTAAA ACCGGGCAAA GTTCCAGGTG	360
TTGGTCTGCC GGGCGTATAC CCGGTGGTG TTCTGCCGGG CGCGCGTTTC CCAGGTGTTG	420
GTGTACTGCC GGGCGTTCCG ACCGGTGCAAG GTGTTAAACC GAAGGCACCA GGTGTAGGCG	480
GCGCGTTCCG GGGTATCCCG GGTGTTGGCC CGTTCGGTGG TCCGCAGCCA GGC GTTCCGC	540
TGGGTTACCC GATCAAAGCG CCGAAGCTTC CAGGTGGCTA CGGTCTGCCG TACACCACCG	600
GTAAACTGCC GTACGGCTAC GGTCCGGGTG GCGTAGCAGG TGCTGCGGGT AAAGCAGGCT	660
ACCCAACCGG TACTGGTGTT GGTCCGCAAG CTGCTGCGGC AGCTGCGGCG AAGGCAGCAG	720
CAAAATTCGG CGCGGGTGCA GCGGGTGTTT TGCCGGGCGT AGGTGGTGCT GGC GTTCCGG	780
GTGTTCCAGG TCGATCCCG GGCATCGGTG GTATCGCAGG CGTAGGTACT CCGGCGGCCG	840

- 29 -

CTGCGGCTGC GGCAGCTGCG GCGAAAGCAG CTAAATACGG TCGGGCAGCA GGCTTGSTTC	900
CGGGTGGTCC AGGCTTCGGT CCGGGTGTAG TAGGCGTTCC GGGTGGTGGT GTTCGGGGCG	960
TAGGTGTTC AGGTGCGGGC ATCCCGGTTG TACCGGGTGC AGGTATCCCG GCGCTGCGG	1020
TTCCAGGTGT TGTATCCCCG GAAGCGGCAG CTAAGGCTGC TGCAGAAAGCT GCGAAATACG	1080
GAGCTCGTCC GGGCGTTGGT GTTGGTGGCA TCCCGACCTA CCGTGTAGGT GCAGGCGGTT	1140
TCCCAGGTTT CCGCGTTGGT GTTGGTGGCA TCCCGGGTGT AGCTGGTGTT CCGTCTGTG	1200
GTGGCGTACC GGGTGTGGT GCGTTCCAG GTGTAGGTAT CTCCCCGAA GCGCAGGCAG	1260
CTGCGGCAGC TAAAGCAGCG AAGTACGGCG TTGGTACTCC GGCGGCAGCA GCTGCTAAAG	1320
CAGCGGCTAA AGCAGCGCAG TTCGGACTAG TTCCGGGCGT AGGTGTTGCG CCAGGTGTTG	1380
GCGTAGCACC GGGTGTGGT GTTGCTCCGG GCGTAGGTCT GGCACCGGGT GTTGGCGTTG	1440
CACCAGGTGT AGGTGTTGCG CCGGGCGTTG GTGTAGCACC GGGTATCGGT CCGGTGGCG	1500
TTGCGGCTGC TGCAGAAATCT GCTGCGAAGG TTGCTGCGAA AGCGCAGCTG CGTGCAGCAG	1560
CTGGTCTGGG TCGGGGCATC CCAGGTCTGG GTGTAGGTGT TGGTGTCCG GGCCTGGGTG	1620
TAGGTGCAGG GGTACCGGGC CTGGGTGTTG GTGCAGGCGT TCCGGGTTTC GGTGCTGGCG	1680
CGGACGAAGG TGTACGTCGT TCCCTGTCTC CAGAACTGCG TGAAGGTGAC CCGTCTCTT	1740
CCCAGCACCT GCCGTCTACC CCGTCTCTC CACGTGTTCC GGGCGCGCTG GCTGCTGCGA	1800
AAGCGGCGAA ATACGGTGCA GCGGTTCGGG GTGTACTGGG CCGTCTGGGT GCTCTGGGCG	1860
GTGTTGGTAT CCCGGGCGGT GTTGTAGGTG CAGGCCCAGC TGCAGCTGCT GCTGCGGCAA	1920
AGGCAGCGGC GAAAGCAGCT CAGTTCGGTC TGGTTGGTGC AGCAGGTCTG GCGGTCTGG	1980
GTGTTGGCGG TCTGGGTGTA CCGGGCGTTG GTGGTCTGGG TGGCATCCCG CCGGCGGCGG	2040
CAGCTAAAGC GGCTAAATAC GGTGCAGCAG GTCTGGGTGG CGTCTGGGT GGTGCTGGTC	2100
AGTTCCCACT GGGCGGTGTA GCGGCACGTC CCGGTTTCGG TCTGTCCCCG ATCTTCCCAG	2160
GCGGTGCATG CCTGGGTAAA GCTTGGCGCC GTAAACGTAA ATAATGATAG	2210

- 30 -

(2) INFORMATION FOR SEQ ID NO:2:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 733 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS:
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:2:

```

Ser Met Gly Gly Val Pro Gly Ala Ile Pro Gly Gly Val Pro Gly Gly
1           5           10           15

Val Phe Tyr Pro Gly Ala Gly Leu Gly Ala Leu Gly Gly Gly Ala Leu
          20           25           30

Gly Pro Gly Gly Lys Pro Leu Lys Pro Val Pro Gly Gly Leu Ala Gly
          35           40           45

Ala Gly Leu Gly Ala Gly Leu Gly Ala Phe Pro Ala Val Thr Phe Pro
          50           55           60

Gly Ala Leu Val Pro Gly Gly Val Ala Asp Ala Ala Ala Tyr Lys
65           70           75           80

Ala Ala Lys Ala Gly Ala Gly Leu Gly Gly Val Pro Gly Val Gly Gly
          85           90           95

Leu Gly Val Ser Ala Gly Ala Val Val Pro Gln Pro Gly Ala Gly Val
          100          105          110

Lys Pro Gly Lys Val Pro Gly Val Gly Leu Pro Gly Val Tyr Pro Gly
          115          120          125

Gly Val Leu Pro Gly Ala Arg Phe Pro Gly Val Gly Val Leu Pro Gly
          130          135          140

Val Pro Thr Gly Ala Gly Val Lys Pro Lys Ala Pro Gly Val Gly Gly
145          150          155          160

Ala Phe Ala Gly Ile Pro Gly Val Gly Pro Phe Gly Gly Pro Gln Pro

```

- 31 -

165	170	175
Gly Val Pro Leu Gly Tyr Pro Ile Lys Ala Pro Lys Leu Pro Gly Gly		
180	185	190
Tyr Gly Leu Pro Tyr Thr Thr Gly Lys Leu Pro Tyr Gly Tyr Gly Pro		
195	200	205
Gly Gly Val Ala Gly Ala Ala Gly Lys Ala Gly Tyr Pro Thr Gly Thr		
210	215	220
Gly Val Gly Pro Gln Ala Ala Ala Ala Ala Ala Lys Ala Ala Ala		
225	230	235 240
Lys Phe Gly Ala Gly Ala Ala Gly Val Leu Pro Gly Val Gly Gly Ala		
245	250	255
Gly Val Pro Gly Val Pro Gly Ala Ile Pro Gly Ile Gly Gly Ile Ala		
260	265	270
Gly Val Gly Thr Pro Ala Ala Ala Ala Ala Ala Ala Ala Ala Lys		
275	280	285
Ala Ala Lys Tyr Gly Ala Ala Ala Gly Leu Val Pro Gly Gly Pro Gly		
290	295	300
Phe Gly Pro Gly Val Val Gly Val Pro Gly Ala Gly Val Pro Gly Val		
305	310	315 320
Gly Val Pro Gly Ala Gly Ile Pro Val Val Pro Gly Ala Gly Ile Pro		
325	330	335
Gly Ala Ala Val Pro Gly Val Val Ser Pro Glu Ala Ala Ala Lys Ala		
340	345	350
Ala Ala Lys Ala Ala Lys Tyr Gly Ala Arg Pro Gly Val Gly Val Gly		
355	360	365
Gly Ile Pro Thr Tyr Gly Val Gly Ala Gly Gly Phe Pro Gly Phe Gly		
370	375	380
Val Gly Val Gly Gly Ile Pro Gly Val Ala Gly Val Pro Ser Val Gly		
385	390	395 400
Gly Val Pro Gly Val Gly Gly Val Pro Gly Val Gly Ile Ser Pro Glu		
405	410	415

- 32 -

Ala Gln Ala Ala Ala Ala Lys Ala Ala Lys Tyr Gly Val Gly Thr
 420 425 430

Pro Ala Ala Ala Ala Ala Lys Ala Ala Ala Lys Ala Ala Gln Phe Gly
 435 440 445

Leu Val Pro Gly Val Gly Val Ala Pro Gly Val Gly Val Ala Pro Gly
 450 455 460

Val Gly Val Ala Pro Gly Val Gly Leu Ala Pro Gly Val Gly Val Ala
 465 470 475 480

Pro Gly Val Gly Val Ala Pro Gly Val Gly Val Ala Pro Gly Ile Gly
 485 490 495

Pro Gly Gly Val Ala Ala Ala Ala Lys Ser Ala Ala Lys Val Ala Ala
 500 505 510

Lys Ala Gln Leu Arg Ala Ala Ala Gly Leu Gly Ala Gly Ile Pro Gly
 515 520 525

Leu Gly Val Gly Val Gly Val Pro Gly Leu Gly Val Gly Ala Gly Val
 530 535 540

Pro Gly Leu C Val Gly Ala Gly Val Pro Gly Phe Gly Ala Gly Ala
 545 550 555 560

Asp Glu Gly Val Arg Arg Ser Leu Ser Pro Glu Leu Arg Glu Gly Asp
 565 570 575

Pro Ser Ser Ser Gln His Leu Pro Ser Thr Pro Ser Ser Pro Arg Val
 580 585 590

Pro Gly Ala Leu Ala Ala Ala Lys Ala Ala Lys Tyr Gly Ala Ala Val
 595 600 605

Pro Gly Val Leu Gly Gly Leu Gly Ala Leu Gly Gly Val Gly Ile Pro
 610 615 620

Gly Gly Val Val Gly Ala Gly Pro Ala Ala Ala Ala Ala Ala Lys
 625 630 635 640

Ala Ala Ala Lys Ala Ala Gln Phe Gly Leu Val Gly Ala Ala Gly Leu
 645 650 655

- 33 -

Gly Gly Leu Gly Val Gly Gly Leu Gly Val Pro Gly Val Gly Gly Leu
 650 655 670

Gly Gly Ile Pro Pro Ala Ala Ala Ala Lys Ala Ala Lys Tyr Gly Ala
 675 680 685

Ala Gly Leu Gly Gly Val Leu Gly Gly Ala Gly Gln Phe Pro Leu Gly
 690 695 700

Gly Val Ala Ala Arg Pro Gly Phe Gly Leu Ser Pro Ile Phe Pro Gly
 705 710 715 720

Gly Ala Cys Leu Gly Lys Ala Cys Gly Arg Lys Arg Lys
 725 730

(2) INFORMATION FOR SEQ ID NO:3:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 698 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS:
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:3:

Gly Gly Val Pro Gly Ala Ile Pro Gly Gly Val Pro Gly Gly Val Phe
 1 5 10 15

Tyr Pro Gly Ala Gly Leu Gly Ala Leu Gly Gly Gly Ala Leu Gly Pro
 20 25 30

Gly Gly Lys Pro Leu Lys Pro Val Pro Gly Gly Leu Ala Gly Ala Gly
 35 40 45

Leu Gly Ala Gly Leu Gly Ala Phe Pro Ala Val Thr Phe Pro Gly Ala
 50 55 60

Leu Val Pro Gly Gly Val Ala Asp Ala Ala Ala Tyr Lys Ala Ala
 65 70 75 80

Leu Gly Ala Gly Leu Gly Gly Val Pro Gly Val Gly Gly Leu Gly

- 34 -

85	90	95
Val Ser Ala Gly Ala Val Val Pro Gln Pro Gly Ala Gly Val Lys Pro		
100	105	110
Gly Lys Val Pro Gly Val Gly Leu Pro Gly Val Tyr Pro Gly Gly Val		
115	120	125
Leu Pro Gly Ala Arg Phe Pro Gly Val Gly Val Leu Pro Gly Val Pro		
130	135	140
Thr Gly Ala Gly Val Lys Pro Lys Ala Pro Gly Val Gly Gly Ala Phe		
145	150	155 160
Ala Gly Ile Pro Gly Val Gly Pro Phe Gly Gly Pro Gln Pro Gly Val		
165	170	175
Pro Leu Gly Tyr Pro Ile Lys Ala Pro Lys Leu Pro Gly Gly Tyr Gly		
180	185	190
Leu Pro Tyr Thr Thr Gly Lys Leu Pro Tyr Gly Tyr Gly Pro Gly Gly		
195	200	205
Val Ala Gly Ala Ala Gly Lys Ala Gly Tyr Pro Thr Gly Thr Gly Val		
210	215	220
Gly Pro Gln Ala Ala Ala Ala Ala Ala Ala Lys Ala Ala Ala Lys Phe		
225	230	235 240
Gly Ala Gly Ala Ala Gly Val Leu Pro Gly Val Gly Gly Ala Gly Val		
245	250	255
Pro Gly Val Pro Gly Ala Ile Pro Gly Ile Gly Gly Ile Ala Gly Val		
260	265	270
Gly Thr Pro Ala Ala Ala Ala Ala Ala Ala Ala Ala Lys Ala Ala		
275	280	285
Lys Tyr Gly Ala Ala Ala Gly Leu Val Pro Gly Gly Pro Gly Phe Gly		
290	295	300
Pro Gly Val Val Gly Val Pro Gly Ala Gly Val Pro Gly Val Gly Val		
305	310	315 320
Pro Gly Ala Gly Ile Pro Val Val Pro Gly Ala Gly Ile Pro Gly Ala		
325	330	335

- 35 -

Ala Val Pro Gly Val Val Ser Pro Glu Ala Ala Ala Lys Ala Ala Ala
340 345 350

Lys Ala Ala Lys Tyr Gly Ala Arg Pro Gly Val Gly Val Gly Gly Ile
355 360 365

Pro Thr Tyr Gly Val Gly Ala Gly Gly Phe Pro Gly Phe Gly Val Gly
370 375 380

Val Gly Gly Ile Pro Gly Val Ala Gly Val Pro Ser Val Gly Gly Val
385 390 395 400

Pro Gly Val Gly Gly Val Pro Gly Val Gly Ile Ser Pro Glu Ala Gln
405 410 415

Ala Ala Ala Ala Ala Lys Ala Ala Lys Tyr Gly Val Gly Thr Pro Ala
420 425 430

Ala Ala Ala Ala Lys Ala Ala Ala Lys Ala Ala Gln Phe Gly Leu Val
435 440 445

Pro Gly Val Gly Val Ala Pro Gly Val Gly Val Ala Pro Gly Val Gly
450 455 460

Val Ala Pro Gly Val Gly Leu Ala Pro Gly Val Gly Val Ala Pro Gly
465 470 475 480

Val Gly Val Ala Pro Gly Val Gly Val Ala Pro Gly Ile Gly Pro Gly
485 490 495

Gly Val Ala Ala Ala Ala Lys Ser Ala Ala Lys Val Ala Ala Lys Ala
500 505 510

Gln Leu Arg Ala Ala Ala Gly Leu Gly Ala Gly Ile Pro Gly Leu Gly
515 520 525

Val Gly Val Gly Val Pro Gly Leu Gly Val Gly Ala Gly Val Pro Gly
530 535 540

Leu Gly Val Gly Ala Gly Val Pro Gly Phe Gly Ala Val Pro Gly Ala
545 550 555 560

Leu Ala Ala Ala Lys Ala Ala Lys Tyr Gly Ala Ala Val Pro Gly Val
565 570 575

- 36 -

Leu Gly Gly Leu Gly Ala Leu Gly Gly Val Gly Ile Pro Gly Gly Val
 580 585 590

Val Gly Ala Gly Pro Ala Ala Ala Ala Ala Ala Lys Ala Ala Ala
 595 600 605

Lys Ala Ala Gln Phe Gly Leu Val Gly Ala Ala Gly Leu Gly Gly Leu
 610 615 620

Gly Val Gly Gly Leu Gly Val Pro Gly Val Gly Gly Leu Gly Gly Ile
 625 630 635 640

Pro Pro Ala Ala Ala Ala Lys Ala Ala Lys Tyr Gly Ala Ala Gly Leu
 645 650 655

Gly Gly Val Leu Gly Gly Ala Gly Gln Phe Pro Leu Gly Gly Val Ala
 660 665 670

Ala Arg Pro Gly Phe Gly Leu Ser Pro Ile Phe Pro Gly Gly Ala Cys
 675 680 685

Leu Gly Lys Ala Cys Gly Arg Lys Arg Lys
 690 695

(2) INFORMATION FOR SEQ ID NO:4:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1983 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(iii) HYPOTHETICAL: YES

(iv) ANTI-SENSE: NO

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:4:

ATGGGTGGCG TTCCGGGTGC TGTTCCGGGT GCGTTCCGG GTGGTGTATT CTACCCAGGC 60
 CCGGGGTTGG GTGGGTTGG GGGTGGCGTT GCAGACGCAG CTGCTGCGTA CAAAGCGGCA 120

AAGGCAGGTG	CGGGTCTGGG	CGGGGTACCA	GGTGTGGG	GTCTGGGTGT	ATGTGTGGG	180
GCAGTGTGTG	CGCAGCGGGG	TGCAGGTGTA	AAAGCGGCA	AAGTTCCAGG	TGTGTGTGTG	240
CCGGGGGTAT	ACCGGGGTTT	CGGTGCTGTT	CGGGGGGGGG	GTTCGGCAGG	TGTGTGTGTA	300
CTGCGGGGGG	TTCGACCGG	TGCAGGTGTT	AAAGCGAAGG	CACCAAGGTG	AGGCGGGGGG	360
TTCGCGGGTA	TCCCGGGTGT	TGGCGCGTTC	GGTGTGGGG	AGCCAGGGGT	TCCGCTGGGT	420
TACCCGATCA	AAGCGCCGAA	GCTTCCAGGT	GGTACGGTG	TGCCGTACAG	CACCGGTAAA	480
CTGCGGTACG	GCTACGGTCC	GGGTGGCGTA	GCAGGTGCTG	CGGGTAAAGG	AGGCTACCCA	540
ACCGGTACTG	GTGTGGGTCC	GCAGGTGCTG	GGGGCAGCTG	CGGGCAAGGG	AGCAGCAAAA	600
TTCGGGGGGG	GTGCAGCGGG	TTCGGGTGCT	GTTCGGGGGG	TAGGTGGTGT	TGGGTGTGGG	660
GGTGTTCAG	GTCCGATCCC	GGGCATCGGT	GGTATCGCAG	GGTAGGTAG	TCCGGGGGGT	720
GCTGCGGCTG	CGGCAGCTGC	CGCGAAAGCA	GCTAAATAGG	GTGGGGCAGG	AGGCTGGGT	780
CCGGGTGGTC	CAGGCTTCGG	TCCGGGTGTT	GTAGGGGTTC	CGGGTTTGGG	TGTGTGTGGG	840
GGCGTAGGTG	TTCAGGTGC	GGGCATCCCG	GTGTATCCGG	GTGCAGGTAT	CCCGGGGGGT	900
GCGGGTTTCG	GTGCTGTATC	CCCGGAAGCG	GCAGCTAAGG	CTGCTGGGAA	AGCTGGGAAA	960
TACCGAGCTC	GTCCGGGGGT	TGGTGTGGGT	GGCATCCCGA	CCTACGGTGT	AGGTGCAGGC	1020
GGTTTCCAG	GTTTCGGGGT	TGGTGTGGGT	GGCATCCCGG	GTGTAGCTGG	TGTTCCGTCT	1080
GTTCGTGGCG	TACCGGGTGT	TGGTGGCGTT	CCAGGTGTAG	GTATCTCCCT	GGAAGCGCAG	1140
GCAGCTGCAG	CAGCTAAAGC	AGCGAAGTAC	GGCGTTGGTA	CTCCGGCGGC	AGCAGCTGCT	1200
AAAGCAGCGG	CTAAAGCAGC	GCAGTTCCGA	CTAGTTCCGG	GGTAGGTGT	TGCSCCAGGT	1260
GTTCGGGTAG	CACCGGGTGT	TGGTGTGGCT	CCGGGGGTAG	GTCTGGCACC	GGGTGTGGGC	1320
GTTCGACCAG	GTGTAGGTGT	TGCGCCGGGC	GTTCGTGTAG	CACCGGGTAT	CGGTCCGGGT	1380
GGCGTTGCGG	CTGCTGGGAA	ATCTGCTGCG	AAGGTTGCTG	CGAAAGCGCA	GCTGCGTGCA	1440
GCAGCTGGTC	TGGTGGGGG	CATCCCAAGT	CTGGGTGTAG	GTGTGGGTGT	TCCGGGGCTG	1500

- 38 -

```

GGTGTAGGTC CAGGGGTACC GGGCCTGGGT GTTGSTGCAG GGGTTCCGGG TTTGGGTGGT 1560
GTTCCGGGGG CGCTGGGTGC TCGGAAAGCG GCGAAATACG GTGCTGTTCC GGTGTACTG 1620
GGCGGTCTCG GTGCTCTGGG CGGTGTTGGT ATCCCGGGCG GTGTTGTAGG TGCAGGTCCA 1680
GCTGCAGCTG CTGCTGCGGC AAAGGCACCG GCGAAAGCAG CTCAGTTCCG TCTGGTTGGT 1740
GCAGCAGGTC TGGGCGGTCT GGGTGTTCGC GGTCTGGGTC TACCGGGCGT TGGTGGTCTG 1800
GGTGGCATCC CGCCGGCGGC GGCAGCTAAA GCGGCTAAAT ACGGTGCAGC AGTCTGGGT 1860
GGCGTCTTGG GTGGTGCTGG TCAGTTCCCA CTGGSCGGTG TAGCGGCACG TCCGGGTTTC 1920
GGTCTGTCCC CGATCTTCCC AGGCGGTGCA TGGCTGGGTA AAGCTTCCCG CGGTAAAAGT 1980
AAA 1983

```

(2) INFORMATION FOR SEQ ID NO:5:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 560 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS:
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:5:

```

Met Gly Gly Val Pro Gly Ala Val Pro Gly Gly Val Pro Gly Gly Val
1           5           10           15

Phe Tyr Pro Gly Ala Gly Phe Gly Ala Val Pro Gly Gly Val Ala Asp
                20           25           30

Ala Ala Ala Ala Tyr Lys Ala Ala Lys Ala Gly Ala Gly Leu Gly Gly
                35           40           45

Val Pro Gly Val Gly Gly Leu Gly Val Ser Ala Gly Ala Val Val Pro
50           55           60

```


- 40 -

305	310	315	320
Gly Ala Arg Pro Gly Val Gly Val Gly Gly Ile Pro Thr Tyr Gly Val			
325	330	335	
Gly Ala Gly Phe Phe Pro Gly Phe Gly Val Gly Val Gly Gly Ile Pro			
340	345	350	
Gly Val Ala Gly Val Pro Ser Val Gly Gly Val Pro Gly Val Gly Gly			
355	360	365	
Val Pro Gly Val Gly Ile Ser Pro Glu Ala Gln Ala Ala Ala Ala Ala			
370	375	380	
Lys Ala Ala Lys Tyr Gly Val Gly Thr Pro Ala Ala Ala Ala Ala Lys			
385	390	395	400
Ala Ala Ala Lys Ala Ala Gln Phe Gly Leu Val Pro Gly Val Gly Val			
405	410	415	
Ala Pro Gly Val Gly Val Ala Pro Gly Val Gly Val Ala Pro Gly Val			
420	425	430	
Gly Leu Ala Pro Gly Val Gly Val Ala Pro Gly Val Gly Val Ala Pro			
435	440	445	
Gly Val Gly Val Ala Pro Gly Ile Gly Pro Gly Gly Val Ala Ala Ala			
450	455	460	
Ala Lys Ser Ala Ala Lys Val Ala Ala Lys Ala Gln Leu Arg Ala Ala			
465	470	475	480
Ala Gly Leu Gly Ala Gly Ile Pro Gly Leu Gly Val Gly Val Gly Val			
485	490	495	
Pro Gly Leu Gly Val Gly Ala Gly Val Pro Gly Leu Gly Val Gly Ala			
500	505	510	
Gly Val Pro Gly Phe Gly Ala Val Pro Gly Ala Leu Ala Ala Ala Lys			
515	520	525	
Ala Ala Lys Tyr Gly Ala Val Pro Gly Val Leu Gly Gly Leu Gly Ala			
530	535	540	
Leu Gly Gly Val Gly Ile Pro Gly Gly Val Val Gly Ala Gly Pro Ala			
545	550	555	560

Ala Ala Ala Ala Ala Ala Lys Ala Ala Ala Lys Ala Ala Gln Phe Gly
565 570 575

Leu Val Gly Ala Ala Gly Leu Gly Gly Leu Gly Val Gly Gly Leu Gly
580 585 590

Val Pro Gly Val Gly Gly Leu Gly Gly Ile Pro Pro Ala Ala Ala Ala
595 600 605

Lys Ala Ala Lys Tyr Gly Ala Ala Gly Leu Gly Gly Val Leu Gly Gly
610 615 620

Ala Gly Gln Phe Pro Leu Gly Gly Val Ala Ala Arg Pro Gly Phe Gly
625 630 635 640

Leu Ser Pro Ile Phe Pro Gly Gly Ala Cys Leu Gly Lys Ala Cys Gly
645 650 655

Arg Lys Arg Lys
660

(2) INFORMATION FOR SEQ ID NO:6:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 441 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(iii) HYPOTHETICAL: YES

(iv) ANTI-SENSE: NO

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:6:

TCCGCCATGG GAGGTGTTCC GGGCGCGCTG GCTGCTGCGA AAGCGGCGAA ATACGGTGCA	60
GCGGTTCCGG GTGTACTGGG CGGTCTGGGT GCTCTGGGCG GTGTTGGTAT CCCGGGCGGT	120
GTTGTAGGTG CAGGCCCAGC TGCAGCTGCT GCTGCGGCAA AGGCAGCGGC GAAAGCAGCT	180

- 42 -

```

CAGTTCGGTC TGGTTGGTGC AGCAGGTGTG GCGGTCTGG GTGTTGGCGG TCTGGGTGTA      240
CCGGGCGTTG GTGGTCTGGG TGGCATCCCG CCGGCGGCGG CAGCTAAAGC GGCTAAATAC      300
GGTGCAGCAG GTCTGGGTGG CGTTCTGGGT GGTGCTGGTC AGTTCCCACT GGGCGGTGTA      360
GCGGCACGTC CGGGTTTCGG TCTGTCCCCG ATCTTCCCAG GCGGTGCATG CCTGGGTAAA      420
GCTTGCGGCC GTAAACGTAA A                                                  441

```

(2) INFORMATION FOR SEQ ID NO:7:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 147 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS:
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:7:

```

Ser Ala Met Gly Gly Val Pro Gly Ala Leu Ala Ala Ala Lys Ala Ala
1           5           10           15

Lys Tyr Gly Ala Ala Val Pro Gly Val Leu Gly Gly Leu Gly Ala Leu
20          25          30

Gly Gly Val Gly Ile Pro Gly Gly Val Val Gly Ala Gly Pro Ala Ala
35          40          45

Ala Ala Ala Ala Ala Lys Ala Ala Ala Lys Ala Ala Gln Phe Gly Leu
50          55          60

Val Gly Ala Ala Gly Leu Gly Gly Leu Gly Val Gly Gly Leu Gly Val
65          70          75          80

Pro Gly Val Gly Gly Leu Gly Gly Ile Pro Pro Ala Ala Ala Ala Lys
85          90          95

Ala Ala Lys Tyr Gly Ala Ala Gly Leu Gly Gly Val Leu Gly Gly Ala
100         105         110

```

- 43 -

Gly Gln Phe Pro Leu Gly Gly Val Ala Ala Arg Pro Gly Phe Gly Leu
 115 120 125

Ser Pro Ile Phe Pro Gly Gly Ala Cys Leu Gly Lys Ala Cys Gly Arg
 130 135 140

Lys Arg Lys
 145

(2) INFORMATION FOR SEQ ID NO:8:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 600 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(iii) HYPOTHETICAL: YES

(iv) ANTI-SENSE: NO

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:8:

TCCGCCATGG GAGCTCTGGT AGGCCTGGGC GTACCGGGCC TGGGTGTTGG TGCAGGCGTT	60
CCGGGTTTCG GTGCTGGCGC GGACGAAGGT GTACGTCGTT CCCTGTCTCC AGAACTGCGT	120
GAAGGTGACC CGTCCTCTTC CCAGCACCTG CCGTCTACCC CGTCCTCTCC ACGTGTTCGG	180
GGCGCGCTGG CTGCTGCGAA AGCGGCGAAA TACGGTGCAG CGGTTCGGG TGTACTGGGC	240
GGTCTGGGTG CTCTGGGCGG TGTGGGTATC CCGGGCGGTG TTGTAGGTGC AGGCCCAGCT	300
GCAGCTGCTG CTGCGGCAAA GGCAGCGGCG AAAGCAGCTC AGTTCGGTCT GGTGGTGCA	360
GCAGGTCTGG GCGGTCTGGG TGTGGCGGT CTGGGTGTAC CGGGCGTTGG TGGTCTGGGT	420
GGCATCCCGC CGGCGGCGGC AGCTAAAGCG GCTAAATACG GTGCAGCAGG TCTGGGTGGC	480
GTTCTGGGTG GTGCTGGTCA GTTCCCACTG GGCGGTGTAG CGGCACGTCC GGGTTTCGGT	540

CTGTCCCCGA TCTTCCCAGG CGGTGCATGC CTGGGTAAAG CTTGCGGCCG TAAACGTAAA 600

(2) INFORMATION FOR SEQ ID NO:9:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 200 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS:
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:9:

Ser	Ala	Met	Gly	Ala	Leu	Val	Gly	Leu	Gly	Val	Pro	Gly	Leu	Gly	Val	1	5	10	15
Gly	Ala	Gly	Val	Pro	Gly	Phe	Gly	Ala	Gly	Ala	Asp	Glu	Gly	Val	Arg	20	25	30	
Arg	Ser	Leu	Ser	Pro	Glu	Leu	Arg	Glu	Gly	Asp	Pro	Ser	Ser	Ser	Gln	35	40	45	
His	Leu	Pro	Ser	Thr	Pro	Ser	Ser	Pro	Arg	Val	Pro	Gly	Ala	Leu	Ala	50	55	60	
Ala	Ala	Lys	Ala	Ala	Lys	Tyr	Gly	Ala	Ala	Val	Pro	Gly	Val	Leu	Gly	65	70	75	80
Gly	Leu	Gly	Ala	Leu	Gly	Gly	Val	Gly	Ile	Pro	Gly	Gly	Val	Val	Gly	85	90	95	
Ala	Gly	Pro	Ala	Ala	Ala	Ala	Ala	Ala	Lys	Ala	Ala	Ala	Lys	Ala		100	105	110	
Ala	Gln	Phe	Gly	Leu	Val	Gly	Ala	Ala	Gly	Leu	Gly	Gly	Leu	Gly	Val	115	120	125	
Gly	Gly	Leu	Gly	Val	Pro	Gly	Val	Gly	Gly	Leu	Gly	Gly	Ile	Pro	Pro	130	135	140	

- 45 -

Ala Ala Ala Ala Lys Ala Ala Lys Tyr Gly Ala Ala Gly Leu Gly Gly
 145 150 155 160

Val Leu Gly Gly Ala Gly Gln Phe Pro Leu Gly Gly Val Ala Ala Arg
 165 170 175

Pro Gly Phe Gly Leu Ser Pro Ile Phe Pro Gly Gly Ala Cys Leu Gly
 180 185 190

Lys Ala Cys Gly Arg Lys Arg Lys
 195 200

(2) INFORMATION FOR SEQ ID NO:10:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 60 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS:
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:10:

Gly Ile Pro Pro Ala Ala Ala Lys Ala Ala Lys Tyr Gly Ala Ala
 1 5 10 15

Gly Leu Gly Gly Val Leu Gly Gly Ala Gly Gln Phe Pro Leu Gly Gly
 20 25 30

Val Ala Ala Arg Pro Gly Phe Gly Leu Ser Pro Ile Phe Pro Gly Gly
 35 40 45

Ala Cys Leu Gly Lys Ala Cys Gly Arg Lys Arg Lys
 50 55 60

(2) INFORMATION FOR SEQ ID NO:11:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 47 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS:
- (D) TOPOLOGY: linear

- 45 -

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:11:

Gly	Ala	Ala	Gly	Leu	Gly	Gly	Val	Leu	Gly	Gly	Ala	Gly	Gln	Phe	Pro
1				5					10					15	
Leu	Gly	Gly	Val	Ala	Ala	Arg	Pro	Gly	Phe	Gly	Leu	Ser	Pro	Ile	Phe
			20					25					30		
Pro	Gly	Gly	Ala	Cys	Leu	Gly	Lys	Ala	Cys	Gly	Arg	Lys	Arg	Lys	
			35				40					45			

(2) INFORMATION FOR SEQ ID NO:12:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 34 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS:
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:12:

Gly	Ala	Asp	Glu	Gly	Val	Arg	Arg	Ser	Leu	Ser	Pro	Glu	Leu	Arg	Glu
1				5					10					15	
Gly	Asp	Pro	Ser	Ser	Ser	Gln	His	Leu	Pro	Ser	Thr	Pro	Ser	Ser	Pro
			20					25				30			
Arg	Val														

(2) INFORMATION FOR SEQ ID NO:13:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 34 amino acids
- (B) TYPE: amino acid

- 47 -

(C) STRANDEDNESS:

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:13:

Gly	Ala	Asp	Glu	Gly	Val	Arg	Arg	Ser	Leu	Ser	Pro	Glu	Leu	Arg	Glu
1				5					10				15		

Gly	Asp	Pro	Ser	Ser	Ser	Gln	His	Leu	Pro	Ser	Thr	Pro	Ser	Ser	Pro
			20					25					30		

Arg Phe

(2) INFORMATION FOR SEQ ID NO:14:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 216 amino acids

(B) TYPE: amino acid

(C) STRANDEDNESS:

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:14:

Ala	Ala	Ala	Gly	Leu	Gly	Ala	Gly	Ile	Pro	Gly	Leu	Gly	Val	Gly	Val
1				5				10					15		

Gly	Val	Pro	Gly	Leu	Gly	Val	Gly	Ala	Gly	Val	Pro	Gly	Leu	Gly	Val
			20					25					30		

Gly	Ala	Gly	Val	Pro	Gly	Phe	Gly	Ala	Gly	Ala	Asp	Glu	Gly	Val	Arg
			35					40					45		

Arg	Ser	Leu	Ser	Pro	Glu	Leu	Arg	Glu	Gly	Asp	Pro	Ser	Ser	Ser	Gln
			50				55					60			

- 48 -

His Leu Pro Ser Thr Pro Ser Ser Pro Arg Val Pro Gly Ala Leu Ala
65 70 75 80

Ala Ala Lys Ala Ala Lys Tyr Gly Ala Ala Val Pro Gly Val Leu Gly
85 90 95

Gly Leu Gly Ala Leu Gly Gly Val Gly Ile Pro Gly Gly Val Val Gly
100 105 110

Ala Gly Pro Ala Ala Ala Ala Ala Ala Ala Lys Ala Ala Ala Lys Ala
115 120 125

Ala Gln Phe Gly Leu Val Gly Ala Ala Gly Leu Gly Gly Leu Gly Val
130 135 140

Gly Gly Leu Gly Val Pro Gly Val Gly Gly Leu Gly Gly Ile Pro Pro
145 150 155 160

Ala Ala Ala Ala Lys Ala Ala Lys Tyr Gly Ala Ala Gly Leu Gly Gly
165 170 175

Val Leu Gly Gly Ala Gly Gln Phe Pro Leu Gly Gly Val Ala Ala Arg
180 185 190

Pro Gly Phe Gly Leu Ser Pro Ile Phe Pro Gly Gly Ala Cys Leu Gly
195 200 205

Lys Ala Cys Gly Arg Lys Arg Lys
210 215

(2) INFORMATION FOR SEQ ID NO:15:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 183 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS:
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:15:

Ala Thr Thr Gly Leu Gly Ile Gly Ile Pro Gly Leu Gly Val Gly Val

- 49 -

1	5	10	15
Gly Val Pro Gly Leu Gly Val Gly Ala Gly Val Pro Gly Leu Gly Val			
20	25	30	
Gly Ala Gly Val Pro Gly Phe Gly Ala Val Pro Gly Ala Leu Ala Ala			
35	40	45	
Ala Lys Ala Ala Lys Tyr Gly Ala Ala Val Pro Gly Val Leu Gly Gly			
50	55	60	
Leu Gly Ala Leu Gly Gly Val Gly Ile Pro Gly Gly Val Val Gly Ala			
65	70	75	80
Gly Pro Ala Ala Ala Ala Ala Ala Lys Ala Ala Ala Lys Ala Ala			
85	90	95	
Gln Phe Gly Leu Val Gly Ala Ala Gly Leu Gly Gly Leu Gly Val Gly			
100	105	110	
Gly Leu Gly Val Pro Gly Val Gly Gly Leu Gly Gly Ile Pro Pro Ala			
115	120	125	
Ala Ala Ala Lys Ala Ala Lys Tyr Gly Ala Ala Gly Leu Gly Gly Val			
130	135	140	
Leu Gly Gly Ala Gly Gln Phe Pro Leu Gly Gly Val Ala Ala Arg Pro			
145	150	155	160
Gly Phe Gly Leu Ser Pro Ile Phe Pro Gly Gly Ala Cys Leu Gly Lys			
165	170	175	
Ala Cys Gly Arg Lys Arg Lys			
180			

THE CLAIMS:

1. A human tropoelastin derivative or an amino acid
sequence variant thereof, wherein the derivative or
5 variant has elastin-like properties.

2. A human tropoelastin derivative or an amino acid
sequence variant thereof, wherein the derivative or
variant has macro-molecular binding properties.

10

3. A derivative or variant thereof according to
claim 2 wherein the macro-molecular binding properties
include the ability to bind glycosaminoglycans.

15

4. A human tropoelastin derivative or an amino acid
sequence variant thereof, wherein the derivative or
variant has elastin-like properties and macro-molecular
binding properties.

20

5. A polynucleotide encoding a derivative or
variant thereof of any one of claims 1 to 4.

25

6. A tropoelastin derivative comprising the amino
acid sequence of SHEL~~Δ~~modified, or an amino acid sequence
variant of the derivative comprising the amino acid
sequence of SHEL~~Δ~~modified.

30

7. A tropoelastin derivative according to claim 6
comprising SEQ ID NO: 5.

35

8. A polynucleotide encoding a tropoelastin
derivative, the derivative comprising the amino acid
sequence of SHEL~~Δ~~modified or an amino acid sequence
variant of the derivative comprising the amino acid
sequence of SHEL~~Δ~~modified.

9. A polynucleotide according to claim 8 comprising
SEQ ID NO: 4.

10. A synthetic polynucleotide encoding a
5 tropoelastin derivative, the derivative comprising the
amino acid sequence of SHELδ26A or an amino acid sequence
variant of the derivative comprising the amino acid
sequence of SHELδ26A.

10 11. A synthetic polynucleotide according to claim
10, the polynucleotide comprising the sequence of from
nucleotide position 1 to 1676 contiguous with the sequence
of from nucleotide position 1775 to 2210 of SEQ ID NO: 1.

15 12. An amino acid sequence variant of the derivative
comprising the amino acid sequence of SHELδ26A.

13. An amino acid sequence variant according to
claim 12 comprising SEQ ID NO:3.

20

14. A tropoelastin derivative comprising the amino
acid sequence of SHELgamma, or an amino acid sequence
variant of the derivative comprising the amino acid
sequence of SHELgamma.

25

15. A tropoelastin derivative according to claim 14
comprising SEQ ID NO:9.

16. A polynucleotide encoding a tropoelastin
30 derivative, the derivative comprising the amino acid
sequence of the derivative SHELgamma, or an amino acid
sequence variant of the derivative comprising the amino
acid sequence of SHELgamma.

35 17. A polynucleotide sequence according to claim 16
comprising SEQ ID NO:8.

- 52 -

18. A tropoelastin derivative comprising the amino acid sequence of SHELgamma excluding exon 26A, or an amino acid sequence variant of the derivative comprising the amino acid sequence of SHELgamma excluding exon 26A.

5

19. A tropoelastin derivative according to claim 18 comprising SEQ ID NO:7.

20. A polynucleotide encoding a tropoelastin derivative, the derivative comprising the amino acid sequence of SHELgamma excluding exon 26A or an amino acid sequence variant of the derivative comprising the amino acid sequence of SHELgamma excluding exon 26A.

21. A polynucleotide sequence according to claim 20 comprising SEQ ID NO: 6.

22. A tropoelastin derivative comprising the amino acid sequence of SHEL31-36, or an amino acid sequence variant of the derivative comprising the amino acid sequence of SHEL31-36.

23. A tropoelastin derivative according to claim 22 comprising SEQ ID NO: 10.

25

24. A polynucleotide encoding a tropoelastin derivative, the derivative comprising the amino acid sequence of SHEL31-36 or an amino acid sequence variant of the derivative comprising the amino acid sequence of SHEL31-36.

30

25. A polynucleotide according to claim 24, the polynucleotide comprising the sequence of from nucleotide position 2022 to 2210 of SEQ ID NO: 1.

35

26. A tropoelastin derivative comprising the amino acid sequence of SHEL32-36, or an amino acid sequence variant of the derivative comprising the amino acid

sequence of SHEL32-36.

27. A tropoelastin derivative according to claim 26 comprising SEQ ID NO: 11.

5

28. A polynucleotide encoding a tropoelastin derivative, the derivative comprising the amino acid sequence of SHEL32-36 or an amino acid sequence variant of the derivative comprising the amino acid sequence of SHEL32-36.

10

29. A polynucleotide according to claim 28, the polynucleotide comprising the sequence of from nucleotide position 2061 to 2210 of SEQ ID NO: 1.

15

30. A tropoelastin derivative comprising the amino acid sequence of peptide 26A, or an amino acid sequence variant of the derivative comprising the amino acid sequence of peptide 26A.

20

31. A tropoelastin derivative according to claim 30 comprising SEQ ID NO: 12 or SEQ ID NO: 13.

32. A polynucleotide encoding a tropoelastin derivative, the derivative comprising the amino acid sequence of peptide 26A or an amino acid sequence variant of the derivative comprising the amino acid sequence of peptide 26A.

25

33. A polynucleotide according to claim 32, the polynucleotide comprising the sequence of from nucleotide position 1677 to 1774 of SEQ ID NO: 1.

30

34. A tropoelastin derivative comprising the amino acid sequence of SHEL26-36, or an amino acid sequence variant of the derivative comprising the amino acid sequence of SHEL26-26.

35

- 54 -

35. A tropoelastin derivative according to claim 34 comprising SEQ ID NO: 14.

36. A polynucleotide encoding a tropoelastin derivative, the derivative comprising the amino acid sequence of SHEL26-36 or an amino acid sequence variant of the derivative comprising the amino acid sequence of SHEL26-36.

37. A polynucleotide according to claim 36, the polynucleotide comprising the sequence of from nucleotide position 1554 to 2210 of SEQ ID NO: 1.

38. A tropoelastin derivative comprising the amino acid sequence of SHEL26-26 excluding exon 26A, or an amino acid sequence variant of the derivative comprising the amino acid sequence of SHEL26-26 excluding exon 26A.

39. A tropoelastin derivative according to claim 38 comprising SEQ ID NO: 15.

40. A polynucleotide encoding a tropoelastin derivative, the derivative comprising the amino acid sequence of SHEL26-26 excluding exon 26A or an amino acid sequence variant of the derivative of SHEL26-26 excluding exon 26A.

41. A polynucleotide according to claim 40, the polynucleotide comprising the sequence of from nucleotide position 1554 to 1676 contiguous with the sequence of from nucleotide position 1776 to 2210 of SEQ ID NO: 1.

42. A vector comprising a polynucleotide according to any one of claims 5, 8, 9, 16, 17, 20, 21, 24, 25, 28, 29, 32, 33, 36, 37, 40 or 41, or a synthetic polynucleotide according to claim 10 or 11.

43. The vector according to claim 42 wherein the

- 55 -

polynucleotide or synthetic polynucleotide is operatively linked to a promoter or enhancer regulatory sequence.

44. The vector according to claim 42 or 43 wherein
5 the polynucleotide or synthetic polynucleotide is
operatively linked to a nucleotide sequence, the
nucleotide sequence encoding a further amino acid
sequence.

10 45. A cell containing a vector according to any one
of claims 42 to 44.

46. A method for producing a derivative of
tropoelastin or an amino acid sequence variant of the
15 derivative, the method comprising:

- (a) providing a vector according to any one of
claims 42 to 44;
- (b) introducing the vector into a cell;
- (c) maintaining the cell in conditions suitable
20 for expression of the vector; and
- (d) isolating the tropoelastin derivative or
variant.

47. A tropoelastin derivative or variant produced by
25 the method of claim 46.

48. A transgenic non-human animal containing a
vector according to any one of claims 42 to 44, or a
polynucleotide according to any one of claims 5, 8, 9, 16,
30 17, 20, 21, 24, 25, 28, 29, 32, 33, 36, 37, 40 or 41, or a
synthetic polynucleotide according to claim 10 or 11.

49. A tropoelastin derivative or variant of the
derivative produced by a transgenic animal according to
35 claim 48

50. method for producing a tropoelastin derivative

- 56 -

claims 1-4, 6, 7, 12-15, 18, 19, 22, 23, 26, 27, 30, 31, 34, 35, 38 or 39, the method comprising producing the tropoelastin derivative or variant by solid-phase peptide synthesis.

5

51. A tropoelastin derivative or variant produced by the method of claim 50.

52. A formulation comprising at least one
10 tropoelastin derivative or variant of the derivative according to any one of 1-4, 6, 7, 12-15, 18, 19, 22, 23, 26, 27, 30, 31, 34, 35, 38, 39, 47 or 49, together with a pharmaceutically acceptable carrier or diluent.

15 53. An expression product comprising a tropoelastin derivative or variant of the derivative according to any one of claims 1-4, 6, 7, 12-15, 18, 19, 22, 23, 26, 27, 30, 31, 34, 35, 38, 39, 47 or 49, and a further amino acid sequence.

20

54. An expression product according to claim 53 wherein the tropoelastin derivative comprises the amino acid sequence of peptide 26A, or an amino acid sequence variant of the derivative comprising the amino acid
25 sequence of peptide 26A.

55. A polynucleotide encoding an expression product according to claims 53 or 54.

30 56. A vector comprising the polynucleotide according to claim 55.

57. A cell containing a vector according to claim 56.

35

58. A method for producing an expression product according to claim 52 or 54, the method comprising:

- 57 -

- (b) introducing the vector into a cell;
- (c) maintaining the cell in conditions suitable for expression of the vector; and
- (d) isolating the expression product.

5

59. An expression product produced by the method of claim 58.

60. An transgenic non-human animal containing a
10 vector according to claim 56 or a polynucleotide according to claim 55.

61. An expression product produced by a transgenic animal according to claim 60.

15

62. A formulation comprising at least one expression product according to any of claims 53, 54, 59 or 61, together with a pharmaceutically acceptable carrier or diluent.

20

63. A hybrid molecule comprising a biological polymer wherein the polymer is linked to a tropoelastin derivative comprising the amino acid sequence of peptide 26A or an amino acid sequence variant of the derivative
25 comprising peptide 26A.

64. A hybrid molecule according to claim 63 wherein the biological polymer is a protein.

65. A hybrid molecule according to claim 64 wherein
30 in the protein is selected from the group consisting of cytokines, growth factors and antibodies.

66. A hybrid molecule according to claim 63 wherein the biological polymer is selected from the group
35 consisting of lipids, sugars and nucleic acids.

67. A polynucleotide sequence encoding a hybrid

68. A vector comprising a polynucleotide sequence according to claim 67.

5 69. A cell containing a vector according to claim 68.

70. A method for producing a hybrid molecule according to claim 64, the method comprising:

- 10 (a) providing a vector according to claim 68;
(b) introducing the vector into a cell;
(c) maintaining the cell in conditions suitable for expression of the vector; and
(d) isolating the hybrid molecule.

15

71. A hybrid molecule produced by the method of claim 70.

20 72. A transgenic non-human animal containing a vector according to claim 68 or a polynucleotide according to claim 67.

73. A hybrid molecule produced by a transgenic animal according to claim 72.

25

74. A hybrid molecule comprising a synthetic polymer linked to peptide 26A or a variant of peptide 26A.

30 75. A formulation comprising at least one hybrid molecule according to any of claims 63-65, 71, 73 and 74, together with a pharmaceutically acceptable carrier or diluent.

35 76. A cross linked complex, the complex comprising at least one of the following:

- (i) at least one derivative or variant of the derivative according to any of 1-4, 6, 7, 12-15,

- 59 -

or 49;

(ii) at least expression product according to any of claims 53, 54, 58 or 61; and

5 (iii) at least one hybrid molecule according to any of claims 63-65, 71, 73 or 74.

77. An implant, the implant comprising at least one of the following:

10 (i) at least one derivative or variant of the derivative according to any of 1-4, 6, 7, 12-15, 18, 19, 22, 23, 26, 27, 30, 31, 34, 35, 38, 39, 47 or 49;

(ii) at least expression product according to any of claims 53, 54, 58 or 61; and

15 (iii) at least one hybrid molecule according to any of claims 63-65, 71, 73 or 74.

78. A method of imparting glycosaminoglycan binding activity to a biological polymer comprising the step of
20 linking a tropoelastin derivative comprising the amino acid sequence of peptide 26A, or an amino acid sequence variant of the derivative comprising the amino acid sequence of peptide 26A with the biological polymer.

25 79. A method of deleting glycosaminoglycan binding activity from a biological polymer comprising the step of deleting a tropoelastin derivative comprising the amino acid sequence of peptide 26A, or an amino acid sequence variant of the derivative comprising the amino acid
30 sequence of peptide 26A from the biological polymer.

80. The method of claim 66 or 67 wherein the biological polymer is a protein.

35 81. A formulation comprising a tropoelastin derivative or variant of the derivative and a synthetic or biological polymer

1/19

1 GATCCATGGGTGGCGTTCCGGGTGCTATCCCGGGTGGCGTTCCGGGTGGTGTATTCTACC 60
 GTACCCACCGCAAGGCCACGATAGGGCCACCGCAAGGCCACCACATAAGATGG
 S H G G V P G A I P G G V P G G V F Y P

61 CAGGCGCGGGTCTGGGTGCACTGGGCGGTGGTGGCTGGGCCCGGGTGGTAAACCGCTGA 120
 GTCCGCGCCAGACCCACGTGACCCGCCACACGCGACCCGGGCCCCACCATTTGGCGACT
 G A G L G A L G G G A L G P G G K P L K

121 AACCGGTTCCAGGCGGTCTGGCAGGTGCTGGTCTGGGTGCAGGTCTGGGCGCGTTCCCGG 180
 TTGGCCAAGGTCCGCCAGACCGTCCACGACCCAGACCCACGTCCAGACCCGCGCAAGGGCC
 P V P G G L A G A G L G A G L G A F P A

181 CGGTTACCTTCCCGGGTGTCTGGTTCGGGTGGCGTTGCAGACGCAGCTGCTGCGTACA 240
 GCCAATGGAAGGGCCCCACGAGACCAAGGGCCACCGCAACGTCTGCGTCGACGACGCATGT
 V T F P G A L V P G G V A D A A A A Y K

241 AAGCGGCAAGGCAGGTGCGGGTCTGGGCGGGGTACCAGGTGTTGGCGGTCTGGGTGTAT 300
 TTCGCGTTTTCGTTCCACGCCCCAGACCCGCCCCATGGTCCACAAACCGCCAGACCCACATA
 A A K A G A G L G G V P G V G G L G V S

301 CTGCTGGCGCAGTTGTTCCGCAAGCGGGTGCAGGTGTAAACCGGGCAAGTTCCAGGTG 360
 GACGACCGCGTCAACAAGGCGTCGGCCCCACGTCCACATTTGGCCCCGTTTCAAGGTCCAC
 A G A V V P Q P G A G V K P G K V P G V

361 TTGGTCTGCGGGGCGTATACCGGGTGGTGTCTGCGGGGCGCGCGTTTCCAGGTGTTG 420
 AACCAAGCGGGCGCATATGGGCCCCACCAAGACGGCCCCCGCGCAAGGGTCCACAAC
 G L P G V Y P G G V L P G A R F P G V G

Figure 1(1)

219

421 GTGTACTGCCGGGCGTTCCGACCGGTGCAGGTGTTAAACCGAAGGCACCAGGTGTAGGCG 480
CACATGACGGGCCCGCAAGGCTGGCCACGTCCACAATTGGCTTCCGTGGTCCACATCCGC
V L P G V P T G A G V K P K A P G V G G

481 GCGCGTTCCGGGGTATCCCGGGTGTGGCCCGTTCCGGTGGTCCGCAGCCAGGCGTTCCGC 540
CGCGCAAGCGCCCATAGGGCCCAACCGGGCAAGCCACGAGGCGTCCGGTCCGCAAGGCG
A F A G I P G V G P F G G P Q P G V P L

541 TGGGTTACCCGATCAAAGCGCCCAAGCTTCCAGGTGGCTACGGTCTGCCGTACACCACCG 600
ACCCAATGGGCTAGTTTCCGGGCTTCGAAGGTCCACCGATGCCAGACGGCATGTGGTGGC
G Y P I K A P K L P G G Y G L P Y T T G

601 GTAAACTGCCGTACGGCTACGGTCCGGGTGGCGTAGCAGGTGCTGCCGGTAAAGCAGGCT 660
CATTTGACGGCATGCCGATGCCAGGCCACCGCATCGTCCACGACGCCCATTTCCGTCCGA
K L P Y G Y G P G G V A G A A G K A G Y

661 ACCCAACCGGTACTGGTGTGGTCCGCAGGCTGCTGCCGGCAGCTGCCGGCGAAGGCAGCAG 720
TGGGTTGGCCATGACCACAACCAAGGCGTCCGACGACGCCGTCGACGCCGCTTCCGTGGTC
P T G T G V G P Q A A A A A A A K A A A

721 CAAAATCCGGCGCGGGTCCAGCGGGTGTCTGCCGGGCGTAGGTGGTGTGGCGTTCCGG 780
GTTTTAAGCCCGCGCCACGTCCGCCACAAGACGGCCCGCATCCACCAGACCGCAAGGCC
K F G A G A A G V L P G V G G A G V P G

781 GTGTTCCAGGTGCGATCCCGGGCATCGGTGGTATCGCAAGCGTAGGTACTCCGGCGGGCCG 840
CACAGGTTCCACGCTAGGGCCCGTAGCCACCATAGCGTCCGCATCCATGAGGCCGGCCGGC
V P G A I P G I G I A G V G T P A A A

841 CTGCGGCTGCCGCACTCCGGCGAAGGCAGCTAAATACGGTGCGGCAGCAGGCCTGGTTC 900
GACGCCGACGCGCTCGACGCCGCTTTCGTCAATTATGCCACGCGGTGGTCCGACCAAG
A A A A A A A K A A K Y G A A A G L V P

3/19

901 CGGGTGGTCCAGGCTTCGGTCCGGGTGTTGTAGGCGTTCCGGGTGCTGGTGTTCGGGGCG 960
GCCCAACAGGTCCGAAGCCAGGCCCAACATCCGCAAGGCCACGACCACAAGGCCCGC
G G P G F G P G V V G V P G A G V P G V

961 TAGGTGTTCCAGGTGCGGGCATCCCGGTTGTACCGGGTGACAGGTATCCCGGGCGCTGCGG 1020
ATCCACAAGGTCCACGCCCGTAGGGCCAAACATGGCCCAAGTCCATAGGGCCCGGACGCC
G V P G A G I P V V P G A G I P G A A V

1021 TTCCAGGTGTTGTATCCCCGGAAGCGGCAGCTAAGGCTGCTGCGAAAGCTGCGAAATACG 1080
AAGGTCCACAACATAGGGGGCCTTCGCCGTGATTCGACGACGCTTTCGACGCTTTATGC
P G V V S P E A A A K A A A K A A K Y G

1081 GAGCTCGTCCGGGCGTTGGTGTGGTGGCATCCCGACCTACGGTGTAGGTGCAGGCGGTT 1140
CTCGAGCAGGCCCGCAACCACAACCACCGTAGGGCTGGATGCCACATCCACGTCCGCCAA
A R P G V G V G G I P T Y G V G A G G F

1141 TCCCAGGTTTCGGCGTTGGTGTGGTGGCATCCCGGGTGTAGCTGGTGTTCGGTCTGTTG 1200
AGGGTCCAAAGCCGCAACCACAACCACCGTAGGGCCCAATCGACCAACAGGCAGACAAC
P G F G V G V G G I P G V A G V P S V G

1201 GTGGCGTACCGGGTGTGGTGGCGTTCCAGGTGTAGGTATCTCCCGGAAGCGCAGGCAG 1260
CACCOCATGGCCCAACACCGCAAGTCCACATCCATAGAGGGGGCCTTCGGGTCCGTC
G V P G V G G V P G V G I S P E A Q A A

1261 CTGCGGCAGCTAAAGCAGCGAAGTACGGCGTTGGTACTCCGGCGGCAGCAGCTGCTAAAG 1320
GACGCCGTGCAATTCGTGCGCTTCATGCCGCAACCATGAGGCCCGCGTCTGTGACGATTC
A A A K A A K Y G V G T P A A A A A K A

1321 CAGCGGCTAAAGCAGCGCAGTTCTGACTAGTTCCGGCGGTAGGTGTTGCCCGCAGTGTG 1380
GTCCCGCATTTCTGTCGGTCAAGCCTGATCAAGGCCCGCATCCACAACGGCGTCCACAAC
A A K A A Q F G L V P G V G V A P G V G

4/19

1381 GCGTAGCACCGGGTGTGGTGTGGCTCCGGGGCGTAGGTCTGGCACCGGGTGTGGCGTTG 1440
CGCATCGTGGGCCACAACCACAACGAGGGCCCGCATCCAGACCGTGGCCCAACAACGCAAC
V A P G V G V A P G V G L A P G V G V A

1441 CACCAGGTGTAGGTGTGGCCCGGGCGTTGGTGTAGCACCGGGTATCGGTCCGGGTGGCG 1500
GTGGTCCACATCCACAACGGCGGGCCGCAACCACATCGTGGCCCATAGCCAGGCCACCGC
P G V G V A P G V G V A P G I G P G G V

1501 TTGCGGCTGCTGCGAAATCTGCTGCGAAGGTTGCTGCGAAAGCGCAGCTCGGTGCAGCAG 1560
AACGCCGACGACGCTTTAGACGACGCTTCCACGACGCTTTCGCGTCGACGCACGTCTGC
A A A A K S A A K V A A K A Q L R A A A

1561 CTGGTCTGGGTGCGGGCATCCCAGGTCTGGGTGTAGGTGTGGTGTTCGGGGCCTGGGTG 1620
GACCAGACCCACGCGCGTAGGGTCCAGACCCACATCCACAACCACAAGGCCCGGACCCAC
G L G A G I P G L G V G V G V P G L G V

1621 TAGGTGCAGGGGTACCGGGCCTGGGTGTGGTGCAGGCGTTCCGGGTTTCGGTGCCTGGCG 1680
ATCCACGTCCCCATGGCCCGGACCCACAACCACGTCCGCAAGGCCCAAGCCACGACCGC
G A G V P G L G V G A G V P G F G A G A

1681 CGGACGAAGGTGTACGTGCTTCCCTGTCTCCAGAACTGCGTGAAGGTGACCCGTCTCTT 1740
GCCTGCTTCCACATGCAGCAAGGACAGAGGTCTTGACGCACTTCCACTGGGCAGGAGAA
D E G V R R S L S P E L R E G D P S S S

1741 CCCAGCACCTGCGGTCTACCCCGTCTCTCCACGTGTTCCGGGCGCGCTGGCTGCTGCGA 1800
GGTTCGTGGACGGCAGATGGGGCAGGAGAGGTGCACAAGGCCCGCGGACGACGACGCT
Q H L P S T P S S P R V P G A L A A A K

1801 AAGCGGCAAAATACGGTGCAACCGTTCCCGGTGTACTGGGCGGTCTGGGTGCTCTGCGCG 1860
TTGCGCGCTTTATGCCACGTCCGCAAGGCCCAATGACCCGCCAGACCCACGAGACCCGC
A A K Y G A A V P G V L G G L G A L G G

5/19

1861	GTGTTGGTATCCCGGGCGGTGTTGTAGGTGCAGGCCAGCTGCAGCTGCTGCTGCGGCAA CACACCATAGGGCCCCGCCACAACATCCACGTCCGGGTGACGTCGACGACGACGCGCGTT	1920
	V G I P G G V V G A G P A A A A A A A K	
1921	AGGCAGCGGCGAAAGCAGCTCAGTTGGGTCTGGTTGGTGACGAGGTCTGGGCGGTCTGG TCCGTGCGCGCTTTGCTCGAGTCAAGCCAGACCAACCAGTCTGTCAGACCCGCCAGACC	1980
	A A A K A A Q F G L V G A A G L G G L G	
1981	GTGTTGGCGGTCTGGGTGTACCGGGCGTTGGTGGTCTGGGTGGCATCCCGCCGGCGCGGG CACACCAGCCAGACCCACATGGCCCCGCAACCACAGACCCACCGTAGGGCGGCGCGCGCC	2040
	V G G L G V P G V G G L G G I P P A A A	
2041	CAGCTAAAGCGGCTAAATACGGTGCAGCAGGTCTGGGTGGCGTTCTGGGTGGTCTGGTC GTGATTTCGCGATTATGCCACGTCTGTCAGACCCACCGCAAGACCCACCGACCGAG	2100
	A K A A K Y G A A G L G G V L G G A G Q	
2101	AGTCCCACTGGGCGGTGTAGCGGCACGTCCGGGTTTCGGTCTGTCCCGATCTTCCAG TCAGGGTTCACCCGCCACATCGCGGTGCAGGCCCAAGCCAGACAGGGGCTAGAAGGGTC	2160
	F P L G G V A A R P G F G L S P I F P G	
2161	CCGGTGCATGCTGGGTAAAGCTTCCGGCCGTAAACGTAAATAATGATG CGCCAGTACGACCCATTTCGAAACGCCCGCATTTGCATTATTAATCTCTAG	2210
	G A C L G K A C G R K R K * * *	

Figure 1(5)

```

1  GGVFGAIPGGVPGGVFFPGAGLGAALGGGALGPGGKPLKPVFGGLAGAGLG 50
  |||
1  GGVFGAIPGGVPGGVFFPGAGLGAALGGGALGPGGKPLKPVFGGLAGAGLG 50
  |||
51  AGLGAFFPAVTFPGALVPGGVADAAAAAYKAAKAGAGLGGVFGVGGGLGVESG 100
  |||
51  AGLGAFFPAVTFPGALVPGGVADAAAAAYKAAKAGAGLGGVFGVGGGLGVESG 100
  |||
101  AVVPQPGAGVFKPKVPGVGLPGVYFGGVLPGARFFGVGVLPVPTGAGVK 150
  |||
101  AVVPQPGAGVFKPKVPGVGLPGVYFGGVLPGARFFGVGVLPVPTGAGVK 150
  |||
151  PKAPGVGGAFAGIPGVGPFGGPQPGVPLQYPIKAPKLPGGTGLPYTTGKL 200
  |||
151  PKAPGVGGAFAGIPGVGPFGGPQPGVPLQYPIKAPKLPGGTGLPYTTGKL 200
  |||
201  FYGYGPGGVAGANGKGYPTGTGVGPGAAAAAAAKAAKFGAGAGVLPG 250
  |||
201  FYGYGPGGVAGANGKGYPTGTGVGPGAAAAAAAKAAKFGAGAGVLPG 250
  |||
251  VGGAGVPGVFGAIPGTGGLAGVGTFAAAAAAAAKAAKYGAAAGLVPGG 300
  |||
251  VGGAGVPGVFGAIPGTGGLAGVGTFAAAAAAAAKAAKYGAAAGLVPGG 300
  |||
301  PGFGPGVVGVPAGVPGVGPAGIPVVPAGIPGAAVPGVVSPEAAAKA 350
  |||
301  PGFGPGVVGVPAGVPGVGPAGIPVVPAGIPGAAVPGVVSPEAAAKA 350
  |||
351  AAKAAKYGARPGVGVGGIPTYGVGAGGFPFGFGVGVGGIPIGVAGVPSVGGV 400
  |||
351  AAKAAKYGARPGVGVGGIPTYGVGAGGFPFGFGVGVGGIPIGVAGVPSVGGV 400
  |||
401  PGVGGVPGVGDISPEAQAAAAAKAAKYGVGTPAAAAAAKAAKAAQFGLVPG 450
  |||
401  PGVGGVPGVGDISPEAQAAAAAKAAKYGVGTPAAAAAAKAAKAAQFGLVPG 450
  |||
451  VGVAPGVGVAPGVGVAPGVGLAPGVGVAPGVGVAPGVGVAPGIGPGGVAA 500
  |||
451  VGVAPGVGVAPGVGVAPGVGLAPGVGVAPGVGVAPGVGVAPGIGPGGVAA 500
  |||
501  AAKAAKVAAKAQLRAAGLGAAGIPGLGVGVGPGGLGVGAGVPGGLGVGAG 550
  |||
501  AAKAAKVAAKAQLRAAGLGAAGIPGLGVGVGPGGLGVGAGVPGGLGVGAG 550
  |||
551  VPGFGAGADEGVESLSPELREOPSSQILPSTPESPVPFGALAAAKAA 600
  |||
551  VPGFGA.....VPGALAAAKAA 567
  |||
601  KFGAAVPGVLGGGLGALGGVGIPIGVVGAGFAAAAAAAKAAKAAQFGLV 650
  |||
601  KFGAAVPGVLGGGLGALGGVGIPIGVVGAGFAAAAAAAKAAKAAQFGLV 617
  |||
651  AAGLGGGLVGGGLVPGVGGGLGGTPAAAAAAKAAKAGAGLGGVGGAGFP 700
  |||
651  AAGLGGGLVGGGLVPGVGGGLGGTPAAAAAAKAAKAGAGLGGVGGAGFP 667
  |||
701  LGGVAAKPGGLAPITFGGACLAACCHIK 731
  |||
668  LGGVAAKPGGLAPITFGGACLAACCHIK 698
  |||

```

Figure 2(1)

```

1  ATGGGTGGCGTTCCGGGTGCTGTTCCGGGTGGCGTTCCGGGTGGTGTATT 50
  |||
1  MetGlyGlyValProGlyAlaValProGlyGlyValProGlyGlyValPh 17
51  CTACCCAGGCGCGGTTTCCGGTCTGTTCCGGGTGGCGTTGCAGACGCAG 100
  |||
18  eTyrProGlyAlaGlyPheGlyAlaValProGlyGlyValAlaAspAla 34
101  CTGCTGCTACAAAGCGGCAAGGCAGGTGCGGGTCTGGGGGGGGTACCA 150
  |||
35  laAlaAlaTyrIysAlaAlaIysAlaGlyAlaGlyLeuGlyGlyValPro 50
151  GGTGTGGGGTCTGGGTGCTCTGCTGGCGCGTGTGTTCCGCAGCGCGG 200
  |||
51  GlyValGlyGlyLeuGlyValSerAlaGlyAlaValValProInProG 67
201  TGCAGGTGTAAACCGGGCAAGTTCAGGTGTGGTCTGCGGGCGGTAT 250
  |||
68  yAlaGlyValIysProGlyIysValProGlyValGlyLeuProGlyValT 84
251  ACCCGGGTTCGGTCTGTTCCGGGGCGCGGTTTCCAGGTGTGGTGTAT 300
  |||
85  yrProGlyPheGlyAlaValProGlyAlaArgPheProGlyValGlyVal 100
301  CTGCGGGCGTTCCGACCGGTGCAGGTGTAAACCGAAGGCACCGAGGTGT 350
  |||
101  LeuProGlyValProThrGlyAlaGlyValIysProIysAlaProGlyVa 117
351  AGCGGGCGCGTTCGCGGGTATCCCGGGTGTGGCCCGTTCGGTGGTCCG 400
  |||
118  lGlyGlyAlaPheAlaGlyTleProGlyValGlyProPheGlyGlyProG 134
401  AGCCAGGGGTTCGCTGGGTATCCCGATCAAGCGCGGAAGCTTCAGGT 450
  |||
135  InProGlyValProLeuGlyTyrProIleIysAlaProIysLeuProGly 150
451  GGCACGGTCTCCGTTACACACCGGTAAACTCCCGTACCGCTACGGTCC 500
  |||
151  GlyTyrGlyLeuProTyrThrThrGlyIysLeuProTyrGlyTyrGlyPr 167
501  GGGTGGCGTTCAGGTGCTGCGGGTAAAGCAAGCTACCCACCGGTACTG 550
  |||
168  cGlyGlyValAlaGlyAlaAlaGlyIysAlaGlyTyrProThrGlyThrG 184
551  GTGTGGTTCGCGAGCTCTGCGCGAGCTGCGGGTAAAGCAAGCTACCCAA 600
  |||
185  IyValGlyProInAlaAlaAlaAlaAlaAlaAlaIysAlaAlaAlaIys 200
601  TTCCGGCGGGTCTCAGCGGTTTCCGGTCTGTTCCGGGTGGCGTTGCTG 650
  |||
201  PheGlyAlaGlyAlaAlaGlyPheGlyAlaValProGlyValGlyGlyAl 217
651  TGCCTTCCCGGTATCCAGTTCGCGTCCCGGCTCCCGTGGTGGTCCAG 700
  |||
218  aGlyValProGlyValProGlyAlaIleProGlyTleGlyGlyAlaAla 234
701  GCGTAACTATCCCGCGCGCGCTCCCGCTCCCGGCTCCCGGCTCCCGG 750
  |||
235  IyValGlyThrProAlaAlaAlaAlaAlaAlaAlaAlaAlaIysAla 250

```

Figure 3(1)

```

751 GCTAATACGGTGGGGCAGCAGGCGCTGGTTCGGGTGGTTCAGGCTTCGG 800
    |||
251 AlaIysTyrGlyAlaAlaAlaGlyLeuValProGlyGlyProGlyPheG 267
    |||
801 TCCGGGTGGTGGAGGGGTTCCGGGTTTCGGTCTGGTTCGGGGGGTAGGTG 850
    |||
268 yProGlyValValGlyValProGlyPheGlyAlaValProGlyValGlyV 284
    |||
851 TTCAGGTGGGGGCATCCCGGTTGACCGGGTGCAGGATCCCGGGGCT 900
    |||
285 aLProGlyAlaGlyTleProValValProGlyAlaGlyTleProGlyAla 300
    |||
901 GCGGGTTTCGGTCTGGTATCCCGGAGGGCGCTAGGCTCTGGGAA 950
    |||
301 AlaGlyPheGlyAlaValSerProGluAlaAlaAlaLysAlaAlaAlaL 317
    |||
951 AGCTGGAAATACGGAGCTGGTCCGGGGCGTGGTGGTGGTGGCATCCCGA 1000
    |||
318 aAlaAlaIysTyrGlyAlaArgProGlyValGlyValGlyGlyTlePro 334
    |||
1001 CCTACGGGTGGGTGCAGGCGGTTTCGGGTTTCGGGTTTCGGTGGT 1050
    |||
335 hTyrGlyValGlyAlaGlyGlyPheProGlyPheGlyValGlyValGly 350
    |||
1051 GGCATCCCGGGTGGAGCTGGTTCGGTCTGGTGGTGGTTCACGGGGT 1100
    |||
351 GlyTleProGlyValAlaGlyValProSerValGlyGlyValProGlyVa 367
    |||
1101 TGGTGGCGTTCCAGGTGTAGGTATCTCCCGGAGGGCAGGCTGGGG 1150
    |||
368 lGlyGlyValProGlyValGlyTleSerProGluAlaGluAlaAlaAla 384
    |||
1151 CAGCTAAGCAGCGAGGTACGGCGTTGGTACTCCGGCGGGCAGCAGCTCT 1200
    |||
385 lAlaAlaLysAlaAlaLysTyrGlyValGlyThrProAlaAlaAlaAla 400
    |||
1201 AAGCAGGGGCTAAGCAGGCGCAGTTCCGACTGGTTCGGGGCGTAGGT 1250
    |||
401 lLysAlaAlaAlaLysAlaAlaGlnPheGlyLeuValProGlyValGlyVa 417
    |||
1251 TGGCCAGGTGGTGGCGTACACCGGGTGTGGTCTCTCCGGGGCGAG 1300
    |||
418 lAlaProGlyValGlyValAlaProGlyValGlyValAlaProGlyValG 434
    |||
1301 GTCTGGCACCGGGTGGTGGCGTGCACCGGTGGAGGTGGTGGCGGGG 1350
    |||
435 lLysAlaProGlyValGlyValAlaProGlyValGlyValAlaProGly 450
    |||
1351 GTTGGTGGTACACCGGGTACCGGTCCGGGTGGCGTTCGGGCTCTGGAA 1400
    |||
451 ValGlyValAlaProGlyTleGlyProGlyGlyValAlaAlaAlaAlaL 467
    |||
1401 ATCTCTGGAAAGGTCTGGAAAGCAGCTGGTGGTGGTGGTGGTGGT 1450
    |||
468 sSerAlaAlaLysValAlaAlaLysAlaGlnLeuArgAlaAlaAlaGly 484
    |||
1451 TGGTGGCGGCTCCCGGCGTGGGTGGTGGTGGTGGTGGTGGTGGTGG 1500
    |||
485 eGlyValGlyTleProGlyLeuGlyValGlyValGlyValProGlyLeu 500

```

Figure 3(2)

9 19

```

1501 GGTGTAGGTGTCAGGGGTACCGGGCCCTGGGTGTGGTGCAGGGCGTTCCGGG 1550
    |||
501 GlyValGlyAlaGlyValProGlyLeuGlyValGlyAlaGlyValProG 517
1551 TTTCGGTGCCTGTTCCGGGGCGCGCTGGCTGCTGGGAAGCGCGGAATACG 1600
    |||
518 yPheGlyAlaValProGlyAlaLeuAlaAlaAlaIysAlaAlaIysTyrG 534
1601 GTGCTGTTCCGGGTGTACTGGCGGTCTGGGTGCTCTGGGCGGTGTGGT 1650
    |||
535 lyAlaValProGlyValLeuGlyGlyLeuGlyAlaLeuGlyGlyValGly 550
1651 ATCCCGGGCGGTGTGTAGGTGCAGGCCAGCTGCAGCTGCTGCTGGCGC 1700
    |||
551 IleProGlyGlyValValGlyAlaGlyProAlaAlaAlaAlaAlaAl 567
1701 AAGGCAGCGGGGAAGCAGCTCAGTTCCGTCCTGGTTGGTGCAGCAGTC 1750
    |||
568 aIysAlaAlaAlaIysAlaAlaGlnPheGlyLeuValGlyAlaAlaGlyL 584
1751 TGGCGGTCCTGGGTGTGGCGGTCTGGGTGTACCGGGCGGTGGTGGTCG 1800
    |||
585 euGlyGlyLeuGlyValGlyGlyLeuGlyValProGlyValGlyGlyLeu 600
1801 GGTGGCATTCCCGCCGGCGCGCGCAGCTAAGGGCTAATACGGTGCAGC 1850
    |||
601 GlyGlyIleProProAlaAlaAlaAlaIysAlaAlaIysTyrGlyAlaAl 617
1851 AGGTCTGGGTGGCGCTCTGGGTGGTGTCTGGTCACTTCCACTGGGCGGTG 1900
    |||
618 aGlyLeuGlyGlyValLeuGlyGlyAlaGlyAlnPheProLeuGlyGlyV 634
1901 TAGCGGCAGGTCCGGGTTTCCGTCTGTCCCGATCTTCCAGGGCGGTGCA 1950
    |||
635 aAlaAlaAlaArgProGlyPheGlyLeuSerProIlePheProGlyGlyAla 650
1951 TGCTTGGGTAAAGCTTGGCGCCGTAACGTAA 1983
    |||
651 CysLeuGlyIysAlaCysGlyArgIysArgIys 661

```

Figure 3(3)

10/19

1 ATGGGTGGCGTTCCGGGTGCTGTTCCGGGTGGCGTTCCGGGTGGTGTATT 50
|||||
1 ATGGGTGGCGTTCCGGGTGCTATCCCGGGTGGCGTTCCGGGTGGTGTATT 50
51 CTACCCAGGCGCGGGTTTCGGTGC..... 74
|||||
51 CTACCCAGGCGCGGGTCTGGGTGCACTGGGCGGTGGTGGCTGGGCCCGG 100
:
:
75TGT 77
||
151 GGTGCAGGTCTGGGCGCGTTCCCGGGGTACCTTCCCGGGTGTCTGGT 200
78 TCCGGGTGGCGTTGCAGACGAGCTGCTGGTACAAAGCGGCAAAGGCAG 127
|||||
201 TCCGGGTGGCGTTGCAGACGAGCTGCTGGTACAAAGCGGCAAAGGCAG 250
128 GTGCGGGTCTGGGCGGGGTACCAGGTGTTGGCGGTCTGGGTGTATCTGCT 177
|||||
251 GTGCGGGTCTGGGCGGGGTACCAGGTGTTGGCGGTCTGGGTGTATCTGCT 300
178 GGCGCAGTTGTTCCGCAGCCGGGTGCAGGTGTAAAACCGGGCAAAGTTCC 227
|||||
301 GGCGCAGTTGTTCCGCAGCCGGGTGCAGGTGTAAAACCGGGCAAAGTTCC 350
228 AGGTGTTGGTCTGCCGGGCGTATACCGGGTTTCGGTGCCTGTTCCGGGCG 277
|||||
351 AGGTGTTGGTCTGCCGGGCGTATACCGGGT...GGTGTCTGCCGGGCG 397
278 CGCGTTTCCAGGTGTTGGTGTACTGCGGGCGTTCCGACCGGTGCAGGT 327
|||||
398 CGCGTTTCCAGGTGTTGGTGTACTGCGGGCGTTCCGACCGGTGCAGGT 447
328 GTTAAACCGAAGGCACCAAGGTGTAGGCGGCGGTTCCGGGTATCCCGGG 377
|||||
448 GTTAAACCGAAGGCACCAAGGTGTAGGCGGCGGTTCCGGGTATCCCGGG 497
378 TGTGGCCCGTTCCGGTGGTCCGCAGCCAGGCGTTCCGCTGGGTATCCCGA 427
|||||
498 TGTGGCCCGTTCCGGTGGTCCGCAGCCAGGCGTTCCGCTGGGTATCCCGA 547
428 TCAAAGCGCGAAGCTTCCAGGTGGCTACGGTCTGCGGTACACCAACCGT 477
|||||
548 TCAAAGCGCGAAGCTTCCAGGTGGCTACGGTCTGCGGTACACCAACCGT 597
478 AAACCTGCGGTACGGCTACGGTCCGGGTGGCGTAGCAGGTGCTGCGGTAA 527
|||||
598 AAACCTGCGGTACGGCTACGGTCCGGGTGGCGTAGCAGGTGCTGCGGTAA 647
528 AGCAGGCTACCCAACCGGTACTGGTGTGGTCCGACAGGCTGCTGCGGCAG 577
|||||
648 AGCAGGCTACCCAACCGGTACTGGTGTGGTCCGACAGGCTGCTGCGGCAG 697
578 CTGCGGCGAAGGCAGCAGCAAAATTCGGGCGGGTGCAGCGGGTTTCGGT 627
|||||
698 CTGCGGCGAAGGCAGCAGCAAAATTCGGGCGGGTGCAGCGGGTTTCGGT 741
628 GCTGTTCCGGGCGTACGGTGGTCTGGCGTTCCGGGTGTTCCAGGTGGAT 677
|||
742 GTTCTGCCGGGCGTACGGTGGTCTGGCGTTCCGGGTGTTCCAGGTGGAT 791

11/19

678 CCOGGGCATCGGTGGTATCGCAGGCGTAGGTACTCCGGCGGCCGCTGCGG 727
|||||
792 CCOGGGCATCGGTGGTATCGCAGGCGTAGGTACTCCGGCGGCCGCTGCGG 841
|||||
728 CTGCGGCAGCTGCGGCGAARGCAGCTAAATACGGTGCGGCAGCAGGCCTG 777
|||||
842 CTGCGGCAGCTGCGGCGAARGCAGCTAAATACGGTGCGGCAGCAGGCCTG 891
|||||
778 GTTCGGGGTGGTCCAGGCTTCGGTCCGGGTGTTGTAGGCGTTCCGGTTT 827
|||||
892 GTTCGGGGTGGTCCAGGCTTCGGTCCGGGTGTTGTAGGCGTTCCGGGT.. 939
|||||
828 CGGTGCTGTTCCGGGCGTAGGTGTTCCAGGTGCGGGCATCCCGTTGTAC 877
|||||
940 .GCTGGTGTTCGGGCGTAGGTGTTCCAGGTGCGGGCATCCCGTTGTAC 988
|||||
878 CGGGTGCAAGGTATCCCGGGCGCTGCGGGTTTCGGTGCTGTATCCCGGAA 927
|||||
989 CGGGTGCAAGGTATCCCGGGCGCTGCGGGTTTCAGGTGTATATCCCGGAA 1038
|||||
928 GCGGCAGCTAAGGCTGCTGCGAARGCTGCGAATAAGGAGCTCGTCCGGG 977
|||||
1039 GCGGCAGCTAAGGCTGCTGCGAARGCTGCGAATAAGGAGCTCGTCCGGG 1088
|||||
978 CGTTGGTGTGGTGGCATCCCGACCTACGGTGTAGGTGCAGGCGGTTTCC 1027
|||||
1089 CGTTGGTGTGGTGGCATCCCGACCTACGGTGTAGGTGCAGGCGGTTTCC 1138
|||||
1028 CAGGTTTGGGCGTTGGTGTGGTGGCATCCCGGGTGTAGCTGGTGTTCG 1077
|||||
1139 CAGGTTTGGGCGTTGGTGTGGTGGCATCCCGGGTGTAGCTGGTGTTCG 1188
|||||
1078 TCTGTTGGTGGCGTACCGGGTGTGGTGGCGTTCCAGGTGTAGGTATCTC 1127
|||||
1189 TCTGTTGGTGGCGTACCGGGTGTGGTGGCGTTCCAGGTGTAGGTATCTC 1238
|||||
1128 CCOGGARGCGCAGGCAGCTGCGGCAGCTAARGCAGCGAAGTACGGCGTTG 1177
|||||
1239 CCOGGARGCGCAGGCAGCTGCGGCAGCTAARGCAGCGAAGTACGGCGTTG 1288
|||||
1178 GTACTCCGGCGGCAGCAGCTGCTAARGCAGCGCTAARGCAGCGCAGTTC 1227
|||||
1289 GTACTCCGGCGGCAGCAGCTGCTAARGCAGCGCTAARGCAGCGCAGTTC 1338
|||||
1228 GGACTAGTTCCGGGCGTAGGTGTTGGGCCAGGTGTTGGCGTAGCACCGGG 1277
|||||
1339 GGACTAGTTCCGGGCGTAGGTGTTGGGCCAGGTGTTGGCGTAGCACCGGG 1388
|||||
1278 TGTGGTGTGCTCCGGGCGTAGGTCTGGCACCGGGTGTGGCGTTGCAC 1327
|||||
1389 TGTGGTGTGCTCCGGGCGTAGGTCTGGCACCGGGTGTGGCGTTGCAC 1438
|||||
1328 CAGGTGTAGGTGTTGGGCCGGGCGTTGGTGTAGCACCGGGTATCGGTCCG 1377
|||||
1439 CAGGTGTAGGTGTTGGGCCGGGCGTTGGTGTAGCACCGGGTATCGGTCCG 1488
|||||
1378 GGTGGCGTTCCGGCTGCTGCGAATCTGCTGCGAAGGTGCTGCGAARGC 1427
|||||
1489 GGTGGCGTTCCGGCTGCTGCGAATCTGCTGCGAAGGTGCTGCGAARGC 1538

Figure 4(2)

12/19

```

1428 GCAGCTGCGTGCAGCAGCTGGTCTGGGTGCGGGCATCCAGGTCTGGGTG 1477
      |||
1539 GCAGCTGCGTGCAGCAGCTGGTCTGGGTGCGGGCATCCAGGTCTGGGTG 1588
      |||
1478 TAGGTGTGTGGTGTTCGGGGCCTGGGTGTAGGTGCAGGGGTACCGGGCCTG 1527
      |||
1589 TAGGTGTGTGGTGTTCGGGGCCTGGGTGTAGGTGCAGGGGTACCGGGCCTG 1638
      |||
1528 GGTGTGTGTGCAGGGCGTTCCGGGTTCGGTGC..... 1559
      |||
1639 GGTGTGTGTGCAGGGCGTTCCGGGTTCGGTGCCTGGCGCGACGAAGGTGT 1688
      |||
      .
      .
1560 .....TGTTCCGGGGCGCGCTGGCT 1578
      |||
1739 AGCACCTGCCGTCTACCCCGTCCCTCTCCACGTGTTCCGGGGCGCGCTGGCT 1788
      |||
1579 GCTGCGAAAGCGGCGAATAACGGT...GCTGTTCGGGGTGTACTGGGCGG 1625
      |||
1789 GCTGCGAAAGCGGCGAATAACGGTGCAGCGGTTCGGGGTGTACTGGGCGG 1838
      |||
1626 TCTGGGTGCTCTGGGCGGTGTGGTATCCCGGGCGGTGTGTAGGTGCAG 1675
      |||
1839 TCTGGGTGCTCTGGGCGGTGTGGTATCCCGGGCGGTGTGTAGGTGCAG 1888
      |||
1676 GCCCAGCTGCAGCTGCTGCTGCGGCAAGGCAGCGGCGAAGCAGCTCAG 1725
      |||
1889 GCCCAGCTGCAGCTGCTGCTGCGGCAAGGCAGCGGCGAAGCAGCTCAG 1938
      |||
1726 TTGGGTCTGGTGTGGTGCAGCAGGTCTGGGCGGTCTGGGTGTGGCGGTCT 1775
      |||
1939 TTCGGTCTGGTGTGGTGCAGCAGGTCTGGGCGGTCTGGGTGTGGCGGTCT 1988
      |||
1776 GGGTGTACCGGGCGTGGTGGTCTGGGTGGCATCCCGCGGGCGGGCGAG 1825
      |||
1989 GGGTGTACCGGGCGTGGTGGTCTGGGTGGCATCCCGCGGGCGGGCGAG 2038
      |||
1826 CTAAGCGGCTAATAACGGTGCAGCAGGTCTGGGTGGCGTCTGGGTGGT 1875
      |||
2039 CTAAGCGGCTAATAACGGTGCAGCAGGTCTGGGTGGCGTCTGGGTGGT 2088
      |||
1876 GCTGGTCAAGTTCCTCACTGGGCGGTGTAGCGGCAGTCCGGGTTCGGTCT 1925
      |||
2089 GCTGGTCAAGTTCCTCACTGGGCGGTGTAGCGGCAGTCCGGGTTCGGTCT 2138
      |||
1926 GTCCCGATCTTCCAGGCGGTGCATGCCCTGGGTAAAGCTTGGCGCGTA 1975
      |||
2139 GTCCCGATCTTCCAGGCGGTGCATGCCCTGGGTAAAGCTTGGCGCGTA 2188
      |||
1976 AACGTAAATAATGATAG 1992
      |||
2189 AACGTAAATAATGATAG 2205
      |||

```

Figure 4(3)

[illegible]

Figure 5(1)

720 730
PPGACLCACGRK
: : : : : : : : :
PPGACLCACGRK
650 660

15/19

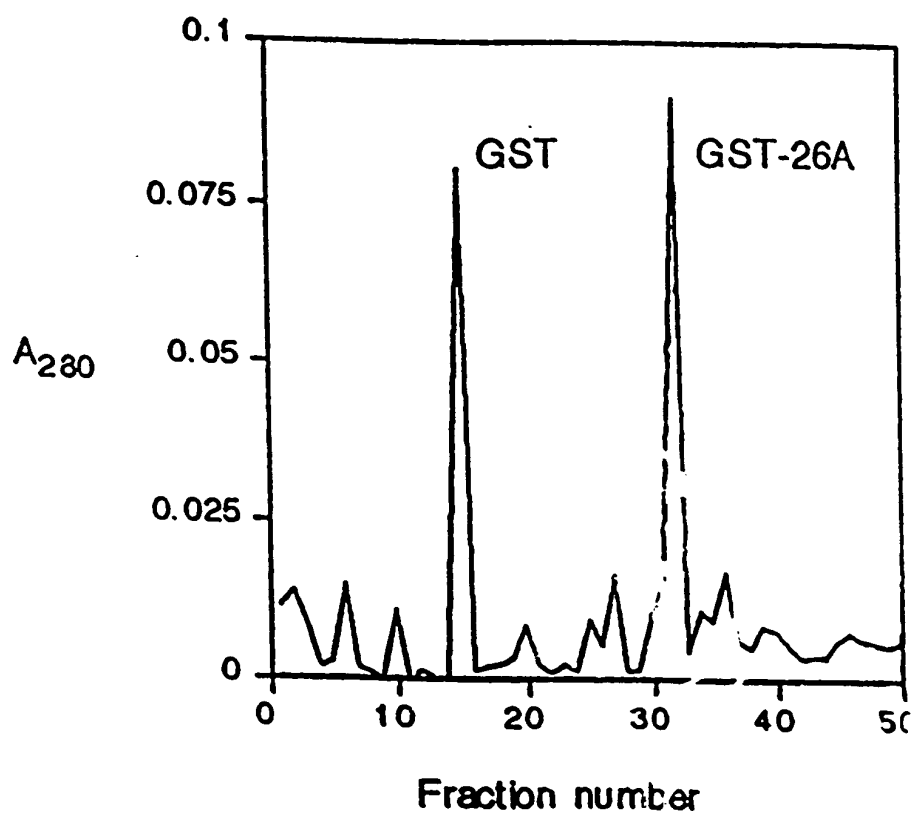


Fig. 6(a)

16/19

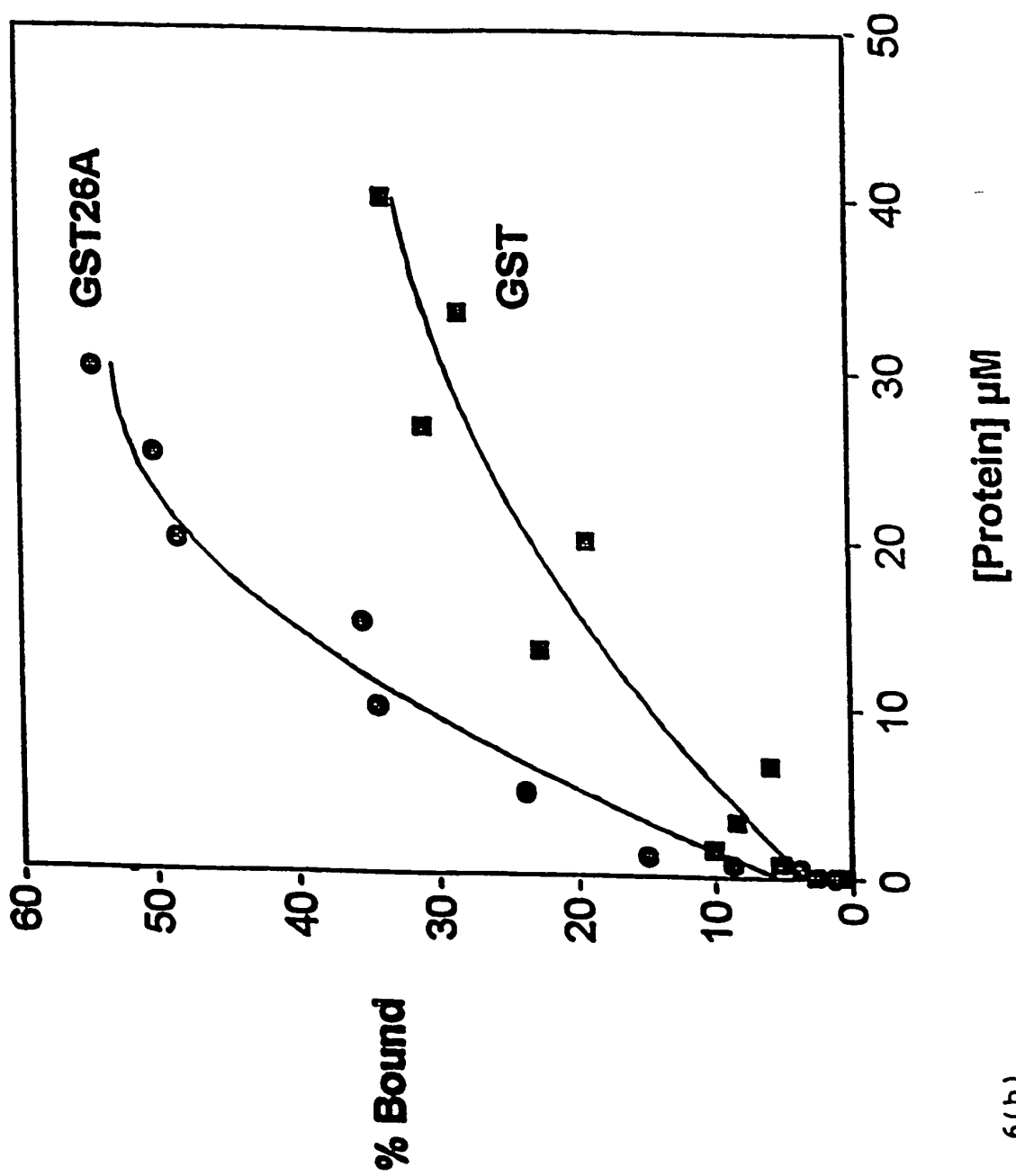


Fig. 6(b)

17/19

948 TCCGCCATGGGAGGTGTTCCGGGCGCGCTGGCTGCTGCGAAAGCGGCGAA 997
|||||
1 SerAlaMetGlyGlyValProGlyAlaLeuAlaAlaAlaLysAlaAlaLy 17

998 ATACGGTGCAGCGGTTCGGGTGTACTGGGCGGTCTGGGTGCTCTGGGCG 1047
|||||
18 sTyrGlyAlaAlaValProGlyValLeuGlyGlyLeuGlyAlaLeuGlyG 34

1048 GTGTTGGTATCCCGGGCGGTGTTGTAGGTGCAGGCCAGCTGCAGCTGCT 1097
|||||
35 lyValGlyIleProGlyGlyValValGlyAlaGlyProAlaAlaAlaAla 50

1098 GCTGCGGCAAAGGCAGCGGCGAAAGCAGCTCAGTTCGGTCTGGTTGGTGC 1147
|||||
51 AlaAlaAlaLysAlaAlaAlaLysAlaAlaGlnPheGlyLeuValGlyAl 67

1148 AGCAGGTCTGGGCGGTCTGGGTGTTGGCGGTCTGGGTGTACCGGGCGTTG 1197
|||||
68 aAlaGlyLeuGlyGlyLeuGlyValGlyGlyLeuGlyValProGlyValG 84

1198 GTGGTCTGGGTGGCATCCCGCCGGCGGCGGCAGCTAAAGCGGCTAAATAC 1247
|||||
85 lyGlyLeuGlyGlyIleProProAlaAlaAlaAlaLysAlaAlaLysTyr 100

1248 GGTGCAGCAGGTCTGGGTGGCGTTCTGGGTGGTGCTGGTCAGTTCCCACT 1297
|||||
101 GlyAlaAlaGlyLeuGlyGlyValLeuGlyGlyAlaGlyGlnPheProLe 117

1298 GGGCGGTGTAGCGGCACGTCCGGGTTTCGGTCTGTCCCGATCTTCCCAG 1347
|||||
118 uGlyGlyValAlaAlaArgProGlyPheGlyLeuSerProIlePheProG 134

1348 GCGGTGCATGCCTGGGTAAAGCTTGCGGCCGTAAACGTAA 1388
|||||
135 lyGlyAlaCysLeuGlyLysAlaCysGlyArgLysArgLys 147

Figure 7

18/19

948 TCCGCCATGGGAGCTCTGGTAGGCCTGGGCGTACCGGGCCTGGGTGTTGG 997
|||||
1 SerAlaMetGlyAlaLeuValGlyLeuGlyValProGlyLeuGlyValGl 17

998 TGCAGGCGTTCCGGGTTTCGGTGCTGGCGCGGACGAAGGTGTACGTCGTT 1047
|||||
18 yAlaGlyValProGlyPheGlyAlaGlyAlaAspGluGlyValArgArgS 34

1048 CCCTGTCTCCAGAACTGCGTGAAGGTGACCCGTCCTCTTCCCAGCACCTG 1097
|||||
35 erLeuSerProGluLeuArgGluGlyAspProSerSerSerGlnHisLeu 50

1098 CCGTCTACCCCGTCCTCTCCACGTGTTCCGGGCGCGCTGGCTGCTGCGAA 1147
|||||
51 ProSerThrProSerSerProArgValProGlyAlaLeuAlaAlaAlaLy 67

1148 AGCGGCGAAATACGGTGCAGCGGTTCCGGGTGTACTGGGCGGTCTGGGTG 1197
|||||
68 sAlaAlaLysTyrGlyAlaAlaValProGlyValLeuGlyGlyLeuGlyA 84

1198 CTCTGGGCGGTGTTGGTATCCCGGGCGGTGTTGTAGGTGCAGGCCAGCT 1247
|||||
85 laLeuGlyGlyValGlyIleProGlyGlyValValGlyAlaGlyProAla 100

Figure 8(1)

19/19

1248 GCAGCTGCTGCTGCGGCAAAGGCAGCGGCGAAAGCAGCTCAGTTCGGTCT 1297
||||||||||||||||||||||||||||||||||||||||||||||||||||||
101 AlaAlaAlaAlaAlaAlaLysAlaAlaAlaLysAlaAlaGlnPheGlyLe 117
1298 GGTGGGTGCAGCAGGTCTGGGCGGTCTGGGTGTTGGCGGTCTGGGTGTAC 1347
||||||||||||||||||||||||||||||||||||||||||||||||||||||
118 uValGlyAlaAlaGlyLeuGlyGlyLeuGlyValGlyGlyLeuGlyValP 134
1348 CGGGCGTTGGTGGTCTGGGTGGCATCCCGCCGGCGGCGGCAGCTAAAGCG 1397
||||||||||||||||||||||||||||||||||||||||||||||||||||||
135 roGlyValGlyGlyLeuGlyGlyIleProProAlaAlaAlaAlaLysAla 150
1398 GCTAAATACGGTGCAGCAGGTCTGGGTGGCGTTCTGGGTGGTGCTGGTCA 1447
||||||||||||||||||||||||||||||||||||||||||||||||||||||
151 AlaLysTyrGlyAlaAlaGlyLeuGlyGlyValLeuGlyGlyAlaGlyGl 167
1448 GTTCCCACTGGGCGGTGTAGCGGCACGTCCGGGTTTCGGTCTGTCCCCGA 1497
||||||||||||||||||||||||||||||||||||||||||||||||||||||
168 nPheProLeuGlyGlyValAlaAlaArgProGlyPheGlyLeuSerProI 184
1498 TCTTCCCAAGCGGTGCATGCCTGGGTAAAGCTTGCAGCGCGTAAACGTAAA 1547
||||||||||||||||||||||||||||||||||||||||||||||||||||||
185 lePheProGlyGlyAlaCysLeuGlyLysAlaCysGlyArgLysArgLys 200

Figure 8(2)